



PR	07-JUN-1995;	95US-0487748.
PR	03-MAR-1995;	95US-0398633.
XX	(MILL-) MILLENNIUM PHARM INC.	
PA	Levinson DA;	
XX	PI	
XX	WI, 1996-433404/43.	
DR	N-PSDB; AAT38265.	
XX		
PT	Genes and their products differentially expressed in T helper cells	
PT	- useful in diagnosis and treatment of immune disorders, e.g.	
PT	multiple sclerosis, asthma, lepromatous leprosy, etc.	
XX		
PS	Claim 10; Fig 17; 218pp; English.	
XX		
CC	The product (AAW01047) of novel murine gene 200 (AAT38265) is a novel	
CC	receptor congly. an extracellular Ig domain. Gene 200 expression is	
CC	many-fold higher in TH1 than in TH2 cell subpopulations.	
CC	Modulation of the 200 gene product may ameliorate a range of	
CC	T-cell-related disorders. Soluble gene 200 products (e.g. fusions	
CC	to immunoglobulins) can be produced that increase the blood	
CC	half-life of the product. Transgenic animals expressing the 200	
CC	gene product are useful models of TH cell subpopulation-related	
CC	disorders. The human homologue (AAW01049) of the gene 200 product	
CC	has also been identified.	
XX		
SQ	Sequence 281 AA;	
	Query Match 100.0%; Score 1468; DB 17; Length 281;	
	Best Local Similarity 100.0%; Pred. No. 2,1e-124;	
	Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 MFSGTLTNCVLLLLQLLLARSLEDEKYVEGKNAYLPCSYTLP TSGTLVPMCMGKGF	60
DB	1 MFSGTLTNCVLLLLQLLLARSLEDEKYVEGKNAYLPCSYTLP TSGTLVPMCMGKGF	60
QY	61 SQCTNELIRTDERNVYTKSSRYOLKGDINKGDSVLIINVTLDHGYCCRIOPFG	120
DB	61 SQCTNELIRTDERNVYTKSSRYOLKGDINKGDSVLIINVTLDHGYCCRIOPFG	120
QY	121 DKLEELKADIRAAKTPAQTAAHGDSTTASPTLTTERNGSETQTLVTLHNNNGTKIST	180
DB	121 DKLEELKADIRAAKTPAQTAAHGDSTTASPTLTTERNGSETQTLVTLHNNNGTKIST	180
QY	181 DEIKDSGETIRTAIHIGVGSAGLTALIIIGVLIIMKWSCKKKLSLSLITANLP	240
DB	181 DEIKDSGETIRTAIHIGVGSAGLTALIIIGVLIIMKWSCKKKLSLSLITANLP	240
QY	241 LANAGAVRIRSEENITYTEENVEYENSNERYCYNNSQOPS	281
DB	241 LANAGAVRIRSEENITYTEENVEYENSNERYCYNNSQOPS	281
RESULT 2		
AA	AA97056	
AC	AA97056 standard; Protein; 281 AA.	
XX	AA97056;	
DT	31-OCT-2000 (first entry)	
DE	Murine T helper cell differentially expressed gene 200 product.	
XX		
KW	T helper cell; differential expression; 200 gene; immunomodulator;	
KW	anti-inflammatory; anti-archival; antibacterial; immunosuppressive;	
KW	thymometric; anti-thyroid; anti-esthatic; anti-allergic; antiviral;	
XX	protozoacide; lymphocyte; modulator; gene therapy.	
CS	Mus sp.	
XX		
Key	Location/Qualifiers	
PT	Peptide 1..20	

FT	Protein	21..281
FT	/label=	Mature protein
FT	Domain	21..192
FT	Domain	/label= Extracellular_domain
FT	Domain	193..214
FT	Domain	/label= Transmembrane_domain
FT	Domain	215..208
FT	/label=	Cytoplasmic_domain
FN	US6084083-A.	
PD	04-JUL-2000.	
XX		
PF	28-MAR-1997;	97US-0829525.
XX		
PR	01-MAR-1996;	96US-0609583.
PR	03-MAR-1995;	95US-0398633.
PR	07-JUN-1995;	95US-0487748.
PA	(MILL-)	MILLENNIUM PHARM INC.
XX		
PI	Levinson DA;	
XX		
DR	WPI: 2000-464385/40.	
DR	N-PSDB; AAA51898.	
XX		
PT	New isolated human 200 gene products or polypeptides, useful for	
PT	treating and diagnosing immune disorders, especially T helper	
PT	lymphocyte-related disorders	
XX		
PS	Example; Fig 17A-D; 107pp; English.	
XX		
CC	Genes which are differentially expressed within and among T helper (TH)	
CC	cells and TH cell subpopulations, e.g. TH0, TH1 and TH2 subpopulations,	
CC	can be used diagnostically or as targets for therapeutic intervention.	
CC	The polypeptides are useful for treating and diagnosing of immune	
CC	disorders, especially T lymphocyte-related disorders. These disorders	
CC	include chronic inflammatory diseases and disorders (e.g. Crohn's	
CC	disease, reactive arthritis, Lyme disease, Hashimoto's thyroiditis or	
CC	Grave's disease), or atopic conditions (e.g. asthma and allergy,	
CC	including allergic rhinitis or food allergies). Also included are	
CC	certain pathogen susceptibilities (e.g. leishmaniasis), and viral (e.g.	
CC	HIV) or bacterial (e.g. tuberculosis or leptomatous leprosy) infections.	
XX		
XX	Sequence	281 AA;
XX		
XX	Query Match	100.0%; Score 1468; DB 21; Length 281;
XX	Best Local Similarity	100.0%; Pred. No. 2,1e-124;
XX	Matches	281; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1	MFGSLTNCVLLLIQLLIARSLDEGKYKEVGKNAYLPSCSYTLPTSGTLVPMCKGKGFPCW 60
DB	1	MFGSLTNCVLLLIQLLIARSLDEGKYKEVGKNAYLPSCSYTLPTSGTLVPMCKGKGFPCW 60
OY	61	SOCTNELRTBERNVTYOKSSRYOLKGDPLNKGDVLIHKNTYLDHGTGCCRIOPGGLAN 120
DB	61	SOCTNELRTBERNVTYOKSSRYOLKGDPLNKGDVLIHKNTYLDHGTGCCRIOPGGLAN 120
OY	121	DKLELKLIDIAKAVTPQAQAHGDSSTASPRTLITERNGSETOTLVTLHNNGTIKISTWA 180
DB	121	DKLELKLIDIAKAVTPQAQAHGDSSTASPRTLITERNGSETOTLVTLHNNGTIKISTWA 180
OY	181	DEIDDSGTTTAAHIGVGSAGLTALIIQVLLIKWVSCKKKKSSSLITLANIPCG 240
DB	181	DEIDDSGTTTAAHIGVGSAGLTALIIQVLLIKWVSCKKKKSSSLITLANIPCG 240
OY	241	LANAGAVIRSEENIYITIEENVYEVENSNEYCYVNSQOPS 281
DB	241	LANAGAVIRSEENIYITIEENVYEVENSNEYCYVNSQOPS 281

ID AAM50225 standard; Protein: 281 AA.  
 AC AAM50225;  
 DT 07-JUN-2002 (first entry)  
 DE Mouse 200 gene product, preferentially expressed in TH1 cells.  
 KM Mouse; 200 gene; T helper; T lymphocyte; T cell; TH1;  
 KM receptor; differential expression; immune disorder; psoriasis;  
 KM multiple sclerosis; insulin-dependent diabetes; antidiabetic;  
 KM antiproliferative; diagnosis; therapy.  
 OS Mus musculus.  
 PN US6288218-B1.  
 PD 11-SEP-2001.  
 PF 25-SEP-1997; 97US-0937339.  
 PR 01-MAR-1996; 96US-0609583.  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 PA (LEVI/) LEVINSON D A.  
 PI Levinson DA;  
 DR WPI; 2001-647189/74.  
 DR N-PSDB; AAI70263, AAI70255.  
 PT Detecting 200 gene expression in a sample, useful for treating and  
 PT diagnosing immune disorders, especially T lymphocyte-related disorders,  
 PT comprising detecting the presence of a 200 gene product or an RNA  
 PT encoding the 200 gene product -  
 PS Claim 2(a); Fig 17A-B; 108pp; English.  
 XX  
 XX The present sequence is that of the protein product of the mouse  
 CC 200 gene (see AAI70263). The 200 gene is preferentially expressed  
 CC in mature, fully differentiated T helper subpopulation TH1 cells  
 CC relative to cells. The gene can be used diagnostically or as  
 CC a target for therapeutic intervention for the treatment of immune  
 CC disorders. A claimed method for detecting 200 gene expression in  
 CC a sample involves detecting the level of a 200 gene product, or an  
 CC RNA encoding it. The gene product detected may be the present  
 CC amino acid sequence, or amino acids 1-20, 1-192, 1-214, 21-192,  
 CC 21-214, 21-281, 293-214, 193-281 or 215-281 of it. Detection of  
 CC the human 200 gene product (see AAM50223) provides a means of  
 CC diagnosing a TH1 cell subpopulation related immune disorder, such  
 CC as multiple sclerosis, psoriasis or insulin-dependent diabetes  
 CC (claimed). In addition to the 200 gene, the invention provides  
 CC other genes that are differentially expressed within and among  
 CC TH cells and TH cell subpopulations and which can be used in  
 CC methods for the diagnosis, prognosis, evaluation and treatment of  
 CC TH cell subpopulation-related disorders, for the identification of  
 CC subjects exhibiting a predisposition to such conditions, for  
 CC monitoring patients undergoing clinical evaluation for the  
 CC treatment of such disorders, and for monitoring the efficacy of  
 CC compounds used in clinical trials. Other immune disorders that can  
 CC be treated/diagnosed include Crohn's disease, reactive arthritis,  
 CC Lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis,  
 CC Grave's disease, contact dermatitis, graft rejection, graft versus  
 CC host disease, sarcoidosis, atopic conditions, asthma, allergy,  
 CC allergic rhinitis, food allergy, eosinophilia, conjunctivitis,  
 CC glomerular nephritis, helminthic infection (e.g. leishmaniasis),  
 CC viral infection (e.g. HIV), and bacterial infection (e.g.  
 CC tuberculosis and lepromatous leprosy).  
 CC  
 CC Sequence 281 AA;  
 SQ Query Match 100.0%; Score 1468; DB 22; Length 281;

Best Local Similarity 100.0%; Pred. No. 2.1e-124;  
 Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MFSGLTNLCVLLQLQLLARSLEDGYKVEGKNAVLPSCSYTLPTSGTLVPMCKGKCPW 60  
 DB 1 MFSGLTNLCVLLQLQLLARSLEDGYKVEGKNAVLPSCSYTLPTSGTLVPMCKGKCPW 60  
 QY 61 SGCNELLRTDERNVYTKSSRYQLKGDNLKNGDVSLIKVNTLDDHGTCCRIOPGLMN 120  
 DB 61 SGCNELLRTDERNVYTKSSRYQLKGDNLKNGDVSLIKVNTLDDHGTCCRIOPGLMN 120  
 QY 121 DKKLELKLIDIKAAKVTPAQTAGDSTTASPRITLTERNGSETOTLVTLHNNNGTKISTWA 180  
 DB 121 DKKLELKLIDIKAAKVTPAQTAGDSTTASPRITLTERNGSETOTLVTLHNNNGTKISTWA 180  
 QY 181 DEIKDGETTIRTAHIGVGSAGITLALLIGVLLKMYSCCKKLSLSLITLANLPFG 240  
 DB 181 DEIKDGETTIRTAHIGVGSAGITLALLIGVLLKMYSCCKKLSLSLITLANLPFG 240  
 QY 241 LANAGAVRIRSEENITYTEENVYEVENSNEYCYVNSQOPS 281  
 DB 241 LANAGAVRIRSEENITYTEENVYEVENSNEYCYVNSQOPS 281  
 RESULT 4  
 ID AAB81511 standard; Protein: 281 AA.  
 AC AAB81511;  
 DT 18-JUN-2001 (first entry)  
 DE Murine TH1 specific 200 gene product.  
 KM Murine; T helper cell; TH cell; TH1; TH2; immunomodulator;  
 KM anti-inflammatory; antiallergic; dermatological; antiviral;  
 KM antibacterial; T helper lymphocyte modulator; gene therapy;  
 KM TH specific gene; 200 gene; immune disorder; inflammation;  
 KM infection.  
 OS Mus sp.  
 PN US6204371-B1.  
 PD 20-MAR-2001.  
 PR 01-MAR-1996; 96US-0609583.  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 PA (MILL-) MILLENNIUM PHARM INC.  
 PI Levinson DA;  
 DR WPI; 2001-272703/28.  
 DR N-PSDB; AAF82609.  
 PT New murine or human 200 genes and their corresponding polypeptides,  
 PT useful for treating or diagnosing immune disorders, especially T helper  
 PT lymphocyte-related disorders, e.g. inflammatory diseases or allergies -  
 PS Claim 5; Fig 17; 109pp; English.  
 XX The present sequence is encoded by murine 200 gene, which is expressed  
 CC at higher levels in T helper (TH)1 cells than in TH2 cells. The  
 CC invention relates to an isolated nucleic acid molecule that comprises  
 CC the full length murine 200 gene or full length human 200 gene nucleotide  
 CC sequence. The nucleic acids are useful for treating or diagnosing immune  
 CC disorders, especially T helper lymphocyte-related disorders, e.g.  
 CC inflammatory diseases (e.g. Crohn's disease), multiple sclerosis, Grave's  
 CC disease, contact dermatitis, psoriasis, asthma and allergies, or certain

CC viral (e.g. HIV) or bacterial (e.g. tuberculosis) infections.  
XX  
SQ Sequence 281 AA;  
Query Match 100.0%; Score 1468; DB 22; Length 281;  
Best Local Similarity 100.0%; Pred. No. 2.1e-124;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MMSGTINCIVLLDQLLLARSLIEDGKYEVGNKAYLPCSYTLPTSGTLVPMCMGKGFPCW 60  
DB 1 MMSGTINCIVLLDQLLLARSLIEDGKYEVGNKAYLPCSYTLPTSGTLVPMCMGKGFPCW 60  
QY 61 SGCETELRTDERNVTYOKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIOPGLMN 120  
DB 61 SGCETELRTDERNVTYOKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIOPGLMN 120  
QY 121 DKLELKLDIRAKVTPAQTAGDSTTASPTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
DB 121 DKLELKLDIRAKVTPAQTAGDSTTASPTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
QY 181 DEIKSGETIRTAIHIGVGSAGLTALIIIGVILIKWYSCKKKLSLITLANLPFG 240  
DB 181 DEIKSGETIRTAIHIGVGSAGLTALIIIGVILIKWYSCKKKLSLITLANLPFG 240  
QY 241 LANAGAVRIRSEENITYTEENYEVENSNERYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENITYTEENYEVENSNERYCYVNSQOPS 281  
RESULT 5  
AAB50696  
ID AAB50696 standard; Protein; 281 AA.  
AC AAB50696;  
XX  
DT 20-MAR-2001 (first entry)  
XX  
DE Mouse 200 gene product amino acid sequence SEQ ID NO:10.  
XX  
KW Treatment; diagnosis; immune disorder; mast cell related disorder;  
KW T-helper lymphocyte-related disorder; ischaemic disorder;  
KW identification; vasodilator; cardiant; antianginal; angina pectoris;  
KW ischaemic renal disease; myocardial ischaemia; myocardial infarction;  
KW cortical infarction; ischaemic injury; kidney transplant.  
XX  
OS Mus musculus.  
XX  
PN WO200073498-A1.  
XX  
PD 07-DEC-2000.  
XX  
PF 31-MAY-2000; 2000WO-US14986.  
XX  
PR 02-JUN-1999; 99US-0324986.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Levinson DA, Lloyd CM, McCarthy SA;  
XX  
DR WPI; 2001-016510/02.  
XX  
DR N-PSDB; AAC90986.  
XX  
PT Ameliorating a symptom of an ischemic disorder or injury in a mammal  
PT e.g. ischemic renal disease or myocardial ischemia, by administering a  
PT 200 gene product (S1), a nucleic acid encoding (S1) or an antibody  
PT directed against (S1).  
XX  
PS Claim 10; Fig 17; 309pp; English.  
XX  
CC The present invention describes a method for ameliorating a symptom of  
CC an ischaemic disorder or injury in a mammal. The method comprises  
CC administering a 200 gene product, a nucleic acid encoding (S1) or an  
CC antibody directed against (S1). The method is useful for treating a

CC symptom of an ischaemic disorder such as ischaemic renal disease or  
CC myocardial ischaemia (such as angina pectoris), myocardial or cortical  
CC infarction. The method is also useful for treating a symptom of an  
CC ischaemic injury occurring due to transplantation of a kidney. The  
CC present sequence is used in the exemplification of the present  
CC invention.  
XX  
SQ Sequence 281 AA;  
Query Match 100.0%; Score 1468; DB 22; Length 281;  
Best Local Similarity 100.0%; Pred. No. 2.1e-124;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MMSGTINCIVLLDQLLLARSLIEDGKYEVGNKAYLPCSYTLPTSGTLVPMCMGKGFPCW 60  
DB 1 MMSGTINCIVLLDQLLLARSLIEDGKYEVGNKAYLPCSYTLPTSGTLVPMCMGKGFPCW 60  
QY 61 SGCETELRTDERNVTYOKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIOPGLMN 120  
DB 61 SGCETELRTDERNVTYOKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIOPGLMN 120  
QY 121 DKLELKLDIRAKVTPAQTAGDSTTASPTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
DB 121 DKLELKLDIRAKVTPAQTAGDSTTASPTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
QY 181 DEIKSGETIRTAIHIGVGSAGLTALIIIGVILIKWYSCKKKLSLITLANLPFG 240  
DB 181 DEIKSGETIRTAIHIGVGSAGLTALIIIGVILIKWYSCKKKLSLITLANLPFG 240  
QY 241 LANAGAVRIRSEENITYTEENYEVENSNERYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENITYTEENYEVENSNERYCYVNSQOPS 281  
RESULT 6  
AAB59162  
ID AAB59162 standard; protein; 281 AA.  
AC AAB59162;  
XX  
DT 21-MAR-2001 (first entry)  
XX  
DE Murine 200 gene protein.  
XX  
KW Cysteine protease; immune disorder; T lymphocyte; Crohn's;  
KW arthritis; diabetes; multiple sclerosis; viral infection; bacterial;  
KW HIV.  
XX  
OS Mus sp..  
XX  
PN US6156887-A.  
XX  
PD 05-DEC-2000.  
XX  
PF 03-OCT-1997; 97US-0939729.  
XX  
PR 01-MAR-1996; 96US-0609583.  
XX  
PR 03-MAR-1995; 95US-0398633.  
XX  
PR 07-JUN-1995; 95US-0487748.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Levinson DA;  
XX  
DR WPI; 2001-101473/11.  
XX  
PT Novel polypeptide exhibiting cysteine protease activity, useful for  
PT treating and diagnosing immune disorders, especially T  
PT lymphocyte-related disorders, e.g. Crohn's disease, multiple sclerosis,  
PT graft versus host disease or allergies -  
XX  
PS Examples; Fig 17; 107pp; English.  
XX



	CC	The present invention relates to a novel cysteine protease. The
	CC	protein of the invention is useful for treating and diagnosing immune
	CC	disorders, especially T lymphocyte-related disorders. In particular,
	CC	the polypeptide is useful for treating or diagnosing T helper (TH) cell
	CC	or TH cell subpopulation-related disorders. These disorders include
	CC	Crohn's disease, reactive arthritis, Lyme disease, insulin-dependent
	CC	diabetes, organ-specific autoimmunity, multiple sclerosis, Hashimoto's
	CC	thyroiditis, Grave's disease, contact dermatitis, psoriasis, graft
	CC	rejection, graft versus host disease, sarcoidosis, atopic (e.g. asthma
	CC	or allergy), eosinophilia, conjunctivitis, glomerular nephritis, or
	CC	helminthic (e.g. leishmaniasis), viral (e.g. HIV (human
	CC	immunodeficiency virus)) or bacterial (e.g. tuberculosis or lepromatous
	CC	leprosy) infections.
SQ	xx	Sequence 281 AA;
Query Match		100.0%; Score 1468; DB 22; Length 281;
Best Local Similarity		100.0%; Pred. No. 2,1e-124;
Matches 281; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
Dd	Db	1 MSGGLTNCVLLILLLOTLARSLDEGKYVEGNAYLPQSYTLPTSGTLVPKMGKGPCPM 60
Qy	Db	1 MSGGLTNCVLLILLLOTLARSLDEGKYVEGNAYLPQSYTLPTSGTLVPKMGKGFCPM 60
Dd	Qy	61 SQCTNELLRTERNVYTKSSRYOLKGDPLNKGDVSLIKNVTLLDHDGYCCRIQPGGLMN 120
Qy	Dd	61 SQCTNELLRTERNVYTKSSRYOLKGDPLNKGDVSLIKNVTLLDHDGYCCRIQPGGLMN 120
Dd	Qy	121 DKLEELKDIAAKATPAQTAAHGSSTASPRLLTERNGSEFQTLVTLHNNGTKISTWA 180
Qy	Dd	121 DKLEELKDIAAKATPAQTAAHGSSTASPRLLTERNGSEFQTLVTLHNNGTKISTWA 180
Dd	Qy	181 DEIDDSGETITAIHIGVVSAGLTALIIIGVLIMKWSSCKKKUSLSLITLANLPFGG 240
Qy	Dd	181 DEIDDSGETITAIHIGVVSAGLTALIIIGVLIMKWSSCKKKUSLSLITLANLPFGG 240
Dd	Qy	241 LANAGAVIRISEENITYITEENVYEVENSNERYCYNNSOOPS 281
Qy	Dd	241 LANAGAVIRISEENITYITEENVYEVENSNERYCYNNSOOPS 281
ID	ABG32767	standard; Protein; 281 AA.
XX	AC	ABG32767;
XX	DT	15-NOV-2002 (first entry)
XX	DE	Mouse Th1 specific 200 protein.
KW	TH1; TH2; antiinflammatory; antidiabetic; antithyroid; atopy;	
KW	antiallergic; antiviral; autoimmune; antiarthritic; dermatological;	
KW	antiapoptotic; nephrotoxic; immunosuppressive; immune response;	
KW	asthma; allergy; allergic rhinitis; viral infection; thyroiditis;	
KW	inflammatory disease; Crohn's disease; arthritis; diabetes; mouse;	
KW	dermatitis; psoriasis; glomerular nephritis; autoimmunity;	
KW	graft rejection.	
OS	Mus sp.	
XX	PN	US6414117-B1.
XX	PD	02-JUL-2002.
XX	PF	12-MAY-1999; 99US-0310367.
XX	PR	01-MAR-1996; 96US-0609583.
XX	PR	28-MAR-1997; 97US-0829525.
XX	PR	03-MAR-1995; 95US-0398633.
XX	PR	07-JUN-1995; 95US-0487748.
XX	PA	(MILL-) MILLENNIUM PHARM INC.

PI	Levinson DA;
PJ	
XX	WPI; 2002-641576/69.
DR	
XX	
PT	Novel polypeptides which are differentially expressed within and among
PR	T helper cells and cell populations, useful for treating immune
PT	disorders, especially T helper cell subpopulation-related disorders -
PS	Example; Fig 17; 109pp; English.
XX	
CC	This invention relates to a novel isolated polypeptide which is a T
CC	helper (Th1) cell subpopulation specific gene product. The invention of
CC	also discloses other Th1 or Th2 specific gene products. The proteins of
CC	the invention may have antiinflammatory, antidiabetic, antithyroid,
CC	antispasmodic, antiallergic, virolic, antirheumatic, dermatological,
CC	antisporadic, nephrotropic and immunosuppressive activities and may be
CC	used as a regulator of the immune response. The proteins of the
CC	invention are also useful to reduce the level of Th2 cell activity for
CC	treating Th1 cell subpopulation-related disorders including atopic
CC	conditions, such as asthma and allergy including allergic rhinitis, the
CC	effects of pathogen, including viral infection, chronic inflammatory
CC	diseases such as Crohn's disease, arthritis, diabetes, thyroiditis,
CC	dermatitis, psoriasis, glomerular nephritis, organ-specific
CC	autoimmunity, graft rejection and graft versus host disease. The
CC	present sequence represents the gene 200 protein isolated from
CC	a Th1/Th2 differential display library, this protein is specifically
CC	expressed in Th1 cells.
CC	
XX	
SQ	Sequence    281 AA;
	Query Match                 100.0%; Score 1468; DB 23; Length 281;
	Best Local Similarity      100.0%; Pred. No. 2.1e-124;
	Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 MFSGTLNCVLLILQLLARSLEPDGYKVEGKNAYLPFCSTLTLPFSGTLPVCWKGGCPW 60
DB	1 MFSGLTNCVLLILQLLARSLEDGYKVEGKNAYLPFCSTLTLPFSGTLPVCWKGGCPW 60
OY	61 SQTCTNELRTDERNAVTVQKSSRYOLKGDNLNKGVSLIIKNVTLDDHGTTCRLOPGLMN 120
DB	61 SQCTNELRTDERNAVTVQKSSRYOLKGDNLNKGVSLIIKNVTLDDHGTTCRLOPGLMN 120
OY	121 DKCLELKLDIAAKAVTPAQTAAHGDSSTAIPRTLTTBENGSETOVLVTLANNGTKISTWA 180
DB	121 DKCLELKLDIAAKAVTPAQTAAHGDSSTAIPRTLTTBENGSETOVLVTLANNGTKISTWA 180
OY	181 DEIDSGEFTTAHIGVGASAGITALLIIQVILLIKKYSCKKKLSSLITLANIPPG 240
DB	181 DEIISGEFTTAHIGVGASAGITALLIIQVILLIKKYSCCKKLSSLITLANIPPG 240
OY	241 LANGAVRIRSEENITYIEENVLEVENSNEYCVNSQOPS 281
DB	241 LANGAVRIRSEENITYIEENVLEVENSNEYCYVNSQOPS 281
RESULT 8	
ABP70434	
ID	ABP70434 standard; Protein; 281 AA.
AC	ABP70434;
XX	
DT	22-APR-2003 (first entry)
XX	
DE	Amino acid sequence of murine TIM-3 BALB/c allele.
XX	
KM	T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KM	TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KM	myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
XX	allergic T cell response; autoimmune disease.
OS	Mus musculus.
XX	

PN W02003002722-A2.  
XX  
XX 09-JAN-2003.  
XX  
XX 01-JUL-2002; 2002WO-US20890.  
XX  
XX 29-JUN-2001; 2001US-302344P.  
XX  
XX (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Mcintire JF, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;  
XX WPI; 2003-210268/20.  
XX DR N-PSDB; AB668328.  
XX  
XX New nucleic acid comprising a mammalian T cell Immunoglobulin domain  
PT and Mucin domain gene sequences, useful for treating cancer or asthma,  
PT allergy, eczema or autoimmune disease -  
XX  
XX Claim 10; Page 75; 94pp; English.  
XX  
XX The present sequence is a murine T cell immunoglobulin domain and  
CC mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1,  
CC TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with  
CC conserved IGV and mucin domains. The locus comprising the TIM family is  
CC genetically associated with immune dysfunction, including asthma. The  
CC TIM gene family is located within a region of human chromosome 5 that  
CC is commonly deleted in malignancies and myelodysplastic syndrome.  
CC Variants of TIM-1 and TIM-3 are associated with susceptibility to  
CC airway hyperactivity and allergic T cell responses, and other variants  
CC associated with protection against these responses. T cells express TIM  
CC proteins, which critically regulate CD4 T cell differentiation. Th1  
CC cells preferentially express TIM-3, while Th2 cells preferentially  
CC express TIM-1. TIM polypeptides and polynucleotides are useful for  
CC treating cancer, asthma, allergies, eczema or autoimmune diseases.  
XX  
XX Sequence 281 AA:  
SQ  
Query Match 100.0%; Score 1468; DB 24; Length 281;  
Best Local Similarity 100.0%; Pred. No. 2.1e-124;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFSGLTNCVLLLOLLARSLSDGKVEGKNAVPCSYTLPTSGLVPMCMGKGFPM 60  
DB 1 MFSGLTNCVLLLOLLARSLSDGKVEGKNAVPCSYTLPTSGLVPMCMGKGFPM 60  
QY 61 SQTCEHLLATDERNVYQKSSRYQKGDINKGVSLLIKNVTLDHGYCCRIQFPGLMN 120  
DB 61 SQTCEHLLATDERNVYQKSSRYQKGDINKGVSLLIKNVTLDHGYCCRIQFPGLMN 120  
QY 121 DKLELKLDIKAAKTPAQTAGDSTTASPRTLTTERNSSEYQTLVTLNNNGTISTWA 180  
DB 121 DKLELKLDIKAAKTPAQTAGDSTTASPRTLTTERNSSEYQTLVTLNNNGTISTWA 180  
QY 181 DEIKDSGETIRTAIHGVASAGLFTALITIGVILKMKYCKKKKLSLITLANLPQGG 240  
DB 181 DEIKDSGETIRTAIHGVASAGLFTALITIGVILKMKYCKKKKLSLITLANLPQGG 240  
QY 241 LANAGAVRIRSEENYTYIEENYEVENSNEYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENYTYIEENYEVENSNEYCYVNSQOPS 281  
RESULT 9  
ID ABG73716 standard; Protein; 281 AA.  
XX ABG73716;  
XX  
XX 09-APR-2003 (first entry)  
XX  
XX Murine TH1-associated 200 protein.  
DE  
XX

KW T-cell receptor; TH; T helper cell; 103 gene; TH2 cell; murine;  
KW TH2 cell marker; TH1 cell; protozoacide; antibacterial; virucide;  
KW immunosuppressive; antiinflammatory; antidiabetic; antidiabetic;  
KW neuroprotective; dermatological; antihypertensive; antidiabetic;  
KW nephrotropic; antiaesthetic antiallergic; CD8 agonist; CD4 agonist;  
KW interleukin agonist; bacterial; viral infection; immune disorder;  
KW Crohn's disease; reactive arthritis; diabetes; multiple sclerosis;  
KW Hashimoto's thyroiditis; Grave's disease; contact dermatitis; psoriasis;  
KW graft rejection; graft versus host disease; asthma; glomerulonephritis;  
KW allergy; gene therapy; TH cell subpopulation.  
XX  
XX Mus musculus.  
XX  
XX US6455685-B1.  
XX  
XX 24-SEP-2002.  
XX  
XX 27-FEB-1998; 98US-0032337.  
XX  
XX 03-MAR-1995; 95US-0398633.  
XX 07-JUN-1995; 95US-0487748.  
XX 01-MAR-1996; 96US-0609583.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Levinson DA;  
XX  
XX WPI; 2003-066247/06.  
XX DR N-PSDB; ABQ77037; ABQ77038.  
XX  
XX Identifying a test compound that binds to a 103 gene product or is a  
PT TH2 cell marker, useful for the identification and therapeutic use of  
PT compounds as treatments of helminthic, bacterial and viral infections,  
PT and immune disorders -  
XX  
XX Example 8; Figure 17A-D; 131pp; English.  
XX  
XX This invention describes a novel method for identifying a test compound  
CC that binds to a 103 gene product or is a TH2 cell marker. The method  
CC comprises contacting a test compound with an immobilised 103 gene  
CC product, removing unbound test compound or separating the complex from  
CC the reaction mixture, and detecting the complex. Identifying a test  
CC compound that binds to a 103 gene product alternatively comprises: (a)  
CC contacting a test compound with a cell engineered to express a 103 gene  
CC product or co-expressing a 103 gene product and a test compound in a  
CC cell; and (b) removing unbound 103 gene product. Identifying a test  
CC compound that is a TH2 cell marker further comprises detecting a complex,  
CC where contacting the test compound with a TH2 and TH1 cell for the test  
CC compound to bind either cell and detecting binding where it indicates the  
CC test compound as a TH2 cell marker. The products described in the  
CC invention have protozoacide, antibacterial, virucide, immunosuppressive,  
CC antiinflammatory, antidiabetic, antidiabetic, neuroprotective,  
CC dermatological, antihypertensive, antipsoriatic, nephrotropic, antiaesthetic;  
CC and antiallergic activity and can act as CD8, CD4 and interleukin  
CC agonists. The methods and compositions of the present invention are  
CC useful for the identification and therapeutic use of compounds as  
CC treatments of helminthic, bacterial and viral infections and immune  
CC disorders such as Crohn's disease, reactive arthritis, diabetes,  
CC multiple sclerosis, Hashimoto's thyroiditis, Grave's disease, contact  
CC dermatitis, psoriasis, graft rejection, graft versus host disease,  
CC asthma, allergy and glomerulonephritis. They can also be used for gene  
CC therapy, for the diagnostic evaluation and prognosis of TH cell  
CC subpopulation-related disorders, identification of subjects exhibiting a  
CC predisposition to such conditions, monitoring undergoing clinical  
CC evaluation and efficacy for the treatment of the disorders. This sequence  
CC represents a T-helper cell associated polypeptide described in the  
CC disclosure of the invention.  
XX  
XX Sequence 281 AA:  
SQ  
Query Match 100.0%; Score 1468; DB 24; Length 281;  
Best Local Similarity 100.0%; Pred. No. 2.1e-124;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSGLTNCVLLLLQLLARSLEDGYKVEGKNAVLPCSYTLPTSGTLVPMCKGKGFPM 60  
 DB 1 MFSGLTNCVLLLLQLLARSLEDGYKVEGKNAVLPCSYTLPTSGTLVPMCKGKGFPM 60  
 QY 61 SQCTNELRTDERNVVYQKSSRYQLKGDVSLIIKNVTLDDHGTCCRIOPPGIMN 120  
 DB 61 SQCTNELRTDERNVVYQKSSRYQLKGDVSLIIKNVTLDDHGTCCRIOPPGIMN 120  
 QY 121 DKLELKLIDIKAAKVTAPQTAHGDSTTASPRTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
 DB 121 DKLELKLIDIKAAKVTAPQTAHGDSTTASPRTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
 QY 181 DEIKDGETTRTAIHIGVGSAGLTALIIIGVLLIKWYSCCKKKLSLSLITLANLPFGG 240  
 DB 181 DEIKDGETTRTAIHIGVGSAGLTALIIIGVLLIKWYSCCKKKLSLSLITLANLPFGG 240  
 QY 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281  
 DB 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281

RESULT 10  
 ABB70435  
 ID ABB70435 standard; Protein; 281 AA.  
 XX  
 AC ABB70435;  
 XX  
 DT 22-APR-2003 (first entry)  
 XX  
 DE Amino acid sequence of murine TIM-3 ES-HBA and DBA/2J allele.

XX  
 KM T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;  
 KM TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;  
 KM myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;  
 KM allergic T cell response; autoimmune disease.  
 XX  
 OS Mus musculus.

XX PN WO2003002722-A2.

XX PD 09-JAN-2003.

XX PF 01-JUL-2002; 2002WO-US20890.

XX PR 29-JUN-2001; 2001US-302344P.

XX PA (STRD ) UNIV LELAND STANFORD JUNIOR.

XX PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;  
 XX DR WPI; 2003-210266/20.  
 XX DR N-PSDB; ABZ68329.

PT New nucleic acid comprising a mammalian T cell immunoglobulin domain  
 PT and mucin domain gene sequences, useful for treating cancer or asthma,  
 PT allergy, eczema or autoimmune disease -  
 XX  
 PS Claim 10; Page 77; 94pp; English.

XX The present sequence is a murine T cell immunoglobulin domain and  
 CC mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1,  
 CC TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with  
 CC conserved IgV and mucin domains. The locus comprising the TIM family is  
 CC genetically associated with immune dysfunction, including asthma. The  
 CC TIM gene family is located within a region of human chromosome 5 that  
 CC is commonly deleted in malignancies and myelodysplastic syndrome.  
 CC Variants of TIM-1 and TIM-3 are associated with susceptibility to  
 CC airway hyperreactivity and allergic T cell responses, and other variants  
 CC associated with protection against these responses. T cells express TIM  
 CC proteins, which critically regulate CD4 T cell differentiation. Th1  
 CC cells preferentially express TIM-3, while Th2 cells preferentially  
 CC express TIM-1. TIM polypeptides and polynucleotides are useful for

CC treating cancer, asthma, allergies, eczema or autoimmune diseases.  
 XX  
 SQ Sequence 281 AA;  
 Query Match 97.2%; Score 1427; DB 24; Length 281;  
 Best Local Similarity 97.5%; Pred. No. 1e-120;  
 Matches 274; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 MFSGLTNCVLLLLQLLARSLEDGYKVEGKNAVLPCSYTLPTSGTLVPMCKGKGFPM 60  
 DB 1 MFSGLTNCVLLLLQLLARSLEDGYKVEGKNAVLPCSYTLPTSGTLVPMCKGKGFPM 60  
 QY 61 SQCTNELRTDERNVVYQKSSRYQLKGDVSLIIKNVTLDDHGTCCRIOPPGIMN 120  
 DB 61 SQCTNELRTDERNVVYQKSSRYQLKGDVSLIIKNVTLDDHGTCCRIOPPGIMN 120  
 QY 121 DKLELKLIDIKAAKVTAPQTAHGDSTTASPRTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
 DB 121 DKLELKLIDIKAAKVTAPQTAHGDSTTASPRTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
 QY 181 DEIKDGETTRTAIHIGVGSAGLTALIIIGVLLIKWYSCCKKKLSLSLITLANLPFGG 240  
 DB 181 DEIKDGETTRTAIHIGVGSAGLTALIIIGVLLIKWYSCCKKKLSLSLITLANLPFGG 240  
 QY 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281  
 DB 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281

RESULT 11  
 AAM01049  
 ID AAM01049 standard; Protein; 301 AA.  
 XX  
 AC AAM01049;  
 XX  
 DT 29-DEC-1996 (first entry)  
 XX  
 DE Product of 200 gene differentially expressed in T helper cells.

XX  
 KM T helper cell; TH cell; T-cell; T-lymphocyte; 200 gene;  
 KM differential expression; immune disorder; multiple sclerosis;  
 KM asthma; lepromatous leprosy; diagnosis; therapy; receptor.  
 XX  
 OS Homo sapiens.

XX FH Key

XX FT Peptide Location/Qualifiers  
 XX FT 1..20 /label= sig\_peptide

XX FT Domain /label= Extracellular\_domain

XX FT Domain /label= Transmembrane\_domain

XX FT Domain /label= Cytoplasmic\_domain

XX FT Domain /label= Cytoplasmic\_domain

XX PN W09627603-A1.

XX PD 12-SEP-1996.

XX PF 01-MAR-1996; 96WO-US02798.

XX PR 07-JUN-1995; 95US-0487748.

XX PR 03-MAR-1995; 95US-0398633.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Levinson DA;

XX DR WPI; 1996-433404/43.

XX DR N-PSDB; AAT38267.

PT Genes and their products differentially expressed in T helper cells  
 PT - useful in diagnosis and treatment of immune disorders, e.g.

PT multiple sclerosis, asthma, lepromatous leprosy, etc.  
PS Claim 10; Fig 24; 218pp; English.  
XX  
XX  
CC The product (AA01049) of novel human gene 200 (AA038267) is a novel  
CC cell surface receptor of the Ig superfamily class. Gene 200  
CC expression is many-fold higher in TH1 than in TH2 subpopulations.  
CC Modulation of the 200 gene product may ameliorate a range of  
CC T-cell-related disorders. Soluble gene 200 products (e.g. fusions  
CC to immunoglobulins) can be produced that increase the blood  
CC half-life of the product. Transgenic animals expressing the 200  
CC gene product are useful models of TH cell subpopulation-related  
CC disorders. The murine homologue (AA01047) of the gene 200 product  
CC has also been identified.  
XX  
SQ Sequence 301 AA;  
Query Match 60.1%; Score 883; DB 17; Length 301;  
Best Local Similarity 63.7%; Pred. No. 1.8e-71;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;  
QY 1 MEGSLTNCVLLQLLARSLEDGYKVEGNAYLPCSYTLPTSGLVPMCGKGFCEW 60  
DB 1 MESHLPFDVCVLLLLLTLSRSEVEYRAVGQNAVLPCTYTPAAPGNLVPVCMGKACPV 60  
QY 61 SOCTNELRTDERNTYOKSSRYOLKGDINKGDVSLIKNTLDDHGTCCRIOPPGLMN 120  
DB 61 FEGGNVLTERTDERDNYW-TSRVWLNGBFRKGDVSLTIENVTLAUSGIYCCRIQIPGIMN 119  
QY 121 DKLELKLIDIKAAKYTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
DB 120 DEKFNVLKIVIRPAKYTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
QY 180 ADE-----IKDGETIRTAIHIGVVSAGLTALIIIVLILKWSCKKKKLSLSL 230  
DB 180 ANELDRSLANDLRDSGATIRIGIYIGAGICAGLALIFGALIFKWSHSEKIONSL 239  
QY 231 ITLANLPBGGLANAGAVIRSEENITYTEENVYEVENSNEYCYVNS-QOPS 281  
DB 240 ISLANLPBGGLANAVABGIRSEENITYTEENVYEVENSNEYCYVNSROOPS 291  
RESULT 12  
AA097058  
ID AA097058 standard; Protein; 301 AA.  
XX  
XX  
AC AA097058;  
XX  
XX  
DT 31-OCT-2000 (first entry)  
XX  
DE Human T helper cell differentially expressed gene 200 product.  
XX  
XX T helper cell; differential expression; 200 gene; immunomodulator;  
KW anti-inflammatory; anti-arthritis; antibacterial; immunosuppressive;  
KW thymomimetic; anti-thyroid; anti-asthmatic; anti-allergic; antiviral;  
KW protozoacide; lymphocyte; modulator; gene therapy.  
XX  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FH Peptide 1..20  
FT /label= Signal\_peptide  
FT Protein 21..301  
FT /label= Mature\_protein  
FT Domain 21..200  
FT /label= Extracellular\_domain  
FT Domain 201..224  
FT /label= Transmembrane\_domain  
FT Domain 225..301  
FT /label= Cytoplasmic\_domain  
XX  
XX US6084083-A.  
XX  
XX

PD 04-JUL-2000.  
XX  
XX  
PF 28-MAR-1997; 97US-0829525.  
XX  
XX 01-MAR-1996; 96US-0609583.  
PR 03-MAR-1995; 95US-0398633.  
PR 07-JUN-1995; 95US-0487748.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
PA  
PI Levinson DA;  
XX  
XX WPI: 2000-464385/40.  
DR N-PSDB; AAA51904, AAA51905.  
DR  
PT New isolated human 200 gene products or polypeptides, useful for  
PT treating and diagnosing immune disorders, especially T helper  
PT lymphocyte-related disorders  
XX  
XX  
PS Claim 1; Fig 24A-D; 107pp; English.  
XX  
XX  
CC Genes which are differentially expressed within and among T helper (TH)  
CC cells and TH cell subpopulations, e.g. TH0, TH1 and TH2 subpopulations,  
CC can be used diagnostically or as targets for therapeutic intervention.  
CC The polypeptides are useful for treating and diagnosing of immune  
CC disorders, especially T lymphocyte-related disorders. These disorders  
CC include chronic inflammatory diseases and disorders (e.g. Crohn's  
CC disease, reactive arthritis, Lyme disease, Hashimoto's thyroiditis or  
CC Grave's disease), or atopic conditions (e.g. asthma and allergy,  
CC including allergic rhinitis or food allergies). Also included are  
CC certain pathogen susceptibilities (e.g. leishmaniasis), and viral (e.g.  
CC HIV) or bacterial (e.g. tuberculosis or lepromatous leprosy) infections.  
XX  
SQ Sequence 301 AA;  
Query Match 60.1%; Score 883; DB 21; Length 301;  
Best Local Similarity 63.7%; Pred. No. 1.8e-71;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;  
QY 1 MEGSLTNCVLLQLLARSLEDGYKVEGNAYLPCSYTLPTSGLVPMCGKGFCEW 60  
DB 1 MESHLPFDVCVLLLLLTLSRSEVEYRAVGQNAVLPCTYTPAAPGNLVPVCMGKACPV 60  
QY 61 SOCTNELRTDERNTYOKSSRYOLKGDINKGDVSLIKNTLDDHGTCCRIOPPGLMN 120  
DB 61 FEGGNVLTERTDERDNYW-TSRVWLNGBFRKGDVSLTIENVTLAUSGIYCCRIQIPGIMN 119  
QY 121 DKLELKLIDIKAAKYTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
DB 120 DEKFNVLKIVIRPAKYTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
QY 180 ADE-----IKDGETIRTAIHIGVVSAGLTALIIIVLILKWSCKKKKLSLSL 230  
DB 180 ANELDRSLANDLRDSGATIRIGIYIGAGICAGLALIFGALIFKWSHSEKIONSL 239  
QY 231 ITLANLPBGGLANAGAVIRSEENITYTEENVYEVENSNEYCYVNS-QOPS 281  
DB 240 ISLANLPBGGLANAVABGIRSEENITYTEENVYEVENSNEYCYVNSROOPS 291  
RESULT 13  
AAM50223  
ID AAM50223 standard; Protein; 301 AA.  
XX  
XX  
AC AAM50223;  
XX  
XX  
DT 07-JAN-2002 (first entry)  
XX  
DE Human 200 gene product, preferentially expressed in TH1 cells.  
XX  
XX Human; 200 gene; T helper; T lymphocyte; T cell; TH1;  
KW receptor; differential expression; immune disorder; psoriasis;  
KW multiple sclerosis; insulin-dependent diabetes; antidiabetic;  
XX  
XX

KW antipsoriatic; diagnosis; therapy.  
 XX Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..20  
 FT /label= Signal\_peptide  
 FT Protein 21..301  
 FT /label= Mature\_protein  
 FT Domain 21..200  
 FT /label= Extracellular\_domain  
 FT Domain 201..224  
 FT /label= Transmembrane\_domain  
 FT Domain 225..301  
 FT /label= Cytoplasmic\_domain  
 XX US6288218-B1.  
 XX  
 XX 11-SEP-2001.  
 XX  
 XX 25-SEP-1997; 97US-0937399.  
 XX  
 XX 01-MAR-1996; 96US-0609583.  
 XX 03-MAR-1995; 95US-0398633.  
 XX 07-JUN-1995; 95US-0487748.  
 XX  
 PA (LEVI/) LEVINSON D A.  
 XX  
 XX Levinson DA;  
 XX  
 XX MPI; 2001-647189/74.  
 XX  
 DR N-PSDB; AAI70254, AAI70281.  
 XX  
 PT Detecting 200 gene expression in a sample, useful for treating and  
 PT diagnosing immune disorders, especially T lymphocyte-related disorders,  
 PT comprising detecting the presence of a 200 gene product or an RNA  
 PT encoding the 200 gene product -  
 XX  
 XX  
 PS Claim 1(a); Fig 24A-B; 108pp; English.  
 XX  
 XX The present sequence is that of the protein product of the human  
 CC 200 gene (see AAI70254). It is a receptor of the Ig superfamily  
 CC class. The 200 gene is preferentially expressed in mature, fully  
 CC differentiated T helper subpopulation TH1 cells relative to  
 CC subpopulation TH2 cells. The gene can be used diagnostically or as  
 CC a target for therapeutic intervention for the treatment of immune  
 CC disorders. A claimed method for diagnosing a TH cell subpopulation  
 CC related immune disorder involves detecting the level of a 200 gene  
 CC product, or an RNA encoding it, so that if the level differs from  
 CC that in a control sample, the disorder is diagnosed. The gene  
 CC product detected may be the present amino acid sequence, or  
 CC amino acids 1-20, 1-200, 1-224, 30-128, 21-200, 21-224, 21-301,  
 CC 201-224, 201-301 or 224-301 of it. The immune disorder is  
 CC especially a TH1 cell subpopulation-related immune disorder, such  
 CC as multiple sclerosis, psoriasis or insulin-dependent diabetes  
 CC (claimed). In addition to the 200 gene, the invention provides  
 CC other genes that are differentially expressed within and among  
 CC TH cells and TH cell subpopulations and which can be used in  
 CC methods for the diagnosis, prognosis, evaluation and treatment of  
 CC TH cell subpopulation-related disorders, for the identification of  
 CC subjects exhibiting a predisposition to such conditions, for  
 CC monitoring patients undergoing clinical evaluation for the  
 CC treatment of such disorders, and for monitoring the efficacy of  
 CC compounds used in clinical trials. Other immune disorders that can  
 CC be treated/diagnosed include Crohn's disease, reactive arthritis,  
 CC Lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis,  
 CC Grave's disease, contact dermatitis, graft rejection, allergic  
 CC host disease, sarcoidosis, atopic conditions, asthma, allergy,  
 CC allergic rhinitis, food allergy, eosinophilia, conjunctivitis,  
 CC glomerular nephritis, helminthic infection (e.g. leishmaniasis),  
 CC viral infection (e.g. HIV), and bacterial infection (e.g.  
 CC tuberculosis and lepromatous leprosy).  
 CC  
 CC

SQL Sequence 301 AA;  
 Query Match 60.1%; Score 883; DB 22; Length 301;  
 Best Local Similarity 63.7%; Pred. No. 1.8e-71;  
 Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;  
 QY 1 MFSGLTNCVLLLLQLLLARSLEDGYKVEGKNAVLPSCYTLPTSGLVPMCKGFCPW 60  
 DB 1 MFSHLPFDCVLLLLLLLRSSSEVEYRAEYGVQNAVLPFCFTPAAPNLVAVCGKAGCFV 60  
 QY 61 SQCTNELLTDERENVYQKSSRYQLKGLDKGVSLIKKVTLLDDHGYTCRQIFGLAN 120  
 DB 61 FECGNVVLRTDERDVYVW-TSRWYLNGBDFRKGDSLTIEVNTLADSGIYCCRQIPIGN 119  
 QY 121 DKLEIKLIDIKAKVTPAQTAFHSDSTTASPRLTTERNG-SEHQTLVTLLHNNGTSTW 179  
 DB 120 DEKFNLKLVKPKKVPAPLQDFPAAPFRMLTTGHGAETQJGSLPDILVLTQSL 179  
 QY 180 ADE-----IKDSGETRTAIHIGVVSAGLTALIGVLLKWSCKKKLSLSL 230  
 DB 180 ANELRDSRLANDRDSGATIRIGIYAGICAGLALALIFGALIFKWSKSKKIQVLSL 239  
 QY 231 ITLANLPQGLANAGVRISEENITTEBNVYEVNSNEYCYVNS-QQPS 281  
 DB 240 ISLANLPQGLANAAVAGIRSEENIYTIENVYEEVEPNEYCYVSSRQOPS 291  
 RESULT 14  
 AAU14409  
 ID AAU14409 standard; Protein; 301 AA.  
 XX  
 AC AAU14409;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human novel protein #280.  
 XX  
 DE Human; novel protein; Antianaemic; osteopathic; antiinflammatory;  
 KW immunomodulatory; cytoskeletal; neuroprotective; vulnerary; nocropic;  
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antiaesthetic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155437-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 25-JAN-2001; 2001WO-US02623.  
 XX  
 PR 25-JAN-2000; 2000US-0491404.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Dymnac RT;  
 XX  
 DR N-PSDB; AAS22714.  
 XX  
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,  
 PT nervous system disorders, and for regenerating bone and cartilage -  
 XX  
 PS Example 4; Page 804-805; 894pp; English.  
 XX  
 CC The invention relates to polynucleotides encoding novel human  
 CC proteins or their active domains. The polypeptides, polynucleotides and  
 CC antibodies raised against the polypeptides are used in a method of  
 CC treatment of a mammal and prevention of disorders caused by the aberrant  
 CC protein expression or activity. The polypeptides can be used as  
 CC molecular weight markers, food supplements, and in antibody production.  
 CC

CC The polypeptides are used to identify compounds which bind to the  
CC polypeptides. Polynucleotides of the invention are used as probes and  
CC primers, for sequencing, for chromosome or gene mapping, in the  
CC production of recombinant proteins, and in generating anti-sense DNA or  
CC RNA and in gene therapy. Polypeptides of the invention can be used to  
CC target drugs to a tumour, in assays to determine biological activity, to  
CC raise antibodies/ elicit an immune response, to determine quantitative  
CC protein levels, as tissue markers, and to isolate receptors or ligands.  
CC Polypeptides of the invention may also be useful in treating platelet  
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
CC the proliferation, differentiation and survival of stem cells, as a  
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
CC fungal infection or from autoimmunity, cancer, allergy, asthma,  
CC graft-versus-host disease, eczema, haemophilia, thrombosis,  
CC anti-inflammatory diseases, nervous system disorders, and infection.  
CC The present sequence represents a protein of the invention.

XX Sequence 301 AA;

Query Match 60.1%; Score 883; DB 22; Length 301;  
Best Local Similarity 63.7%; Pred. No. 1.8e-71;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSGLTNCVLLLLQLLARSLEDGYKVEGNALPFCSTYPTPTSGTLVPMCKWGFCPW 60  
DB 1 MFSHLPFCVLLLLLLLRSEVEYRAVEGQNALPCFTYPAAGNALVPVCMGKACPV 60  
QY 61 SOCTNELRTDERNTYOKSSRYOLKGDVSLIKNTLDDHGTCCRIOPPGILMN 120  
DB 61 FECGVVARTDERDYNW-TSRWMLNGDFRGKDVSLTENVTLADSGIYCCRIQIPGILMN 119  
QY 121 DKLELKIDIAKAVTPAQTAHGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
DB 120 DEKFMLKLVIRPAKVTAPPTLQORDFTAFPRMLTTRGHGPAFTQTLGSLPDLINLQIISTL 179  
QY 160 ADE-----IKSGEITRTAHIGVGSAGLTALIIQVLLIKWSSCKKKLSLSTL 230  
DB 160 ANELRDSRLANDLRSGATIRIGIYGAGICAGLALIFGALIFKWSHSEKIQNLSTL 239  
QY 231 ITLANLPFGGLANAGAVIRSEENIYTIENVEYVENSNEYCYVNS-QQPS 281  
DB 240 ISLANLPBGLANAVABGIRSEENIYTIENVEYVEEENEYCYVNSQQPS 291

RESULT 15

AAB93838  
ID AAB93838 standard; Protein; 301 AA.

XX AAB93838;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:13669.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EPI074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

XX Ota T, Iogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
DR WPI; 2001-318749/34.

PT primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -

PS Claim 8; SEQ ID 13669; 2537bp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH3166 to AAH13628 and  
CC AAH3633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

SQ Sequence 301 AA;

Query Match 60.1%; Score 883; DB 22; Length 301;  
Best Local Similarity 63.7%; Pred. No. 1.8e-71;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSGLTNCVLLLLQLLARSLEDGYKVEGNALPFCSTYPTPTSGTLVPMCKWGFCPW 60  
DB 1 MFSHLPFCVLLLLLLLRSEVEYRAVEGQNALPCFTYPAAGNALVPVCMGKACPV 60  
QY 61 SOCTNELRTDERNTYOKSSRYOLKGDVSLIKNTLDDHGTCCRIOPPGILMN 120  
DB 61 FECGVVARTDERDYNW-TSRWMLNGDFRGKDVSLTENVTLADSGIYCCRIQIPGILMN 119  
QY 121 DKLELKIDIAKAVTPAQTAHGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
DB 120 DEKFMLKLVIRPAKVTAPPTLQORDFTAFPRMLTTRGHGPAFTQTLGSLPDLINLQIISTL 179  
QY 160 ADE-----IKSGEITRTAHIGVGSAGLTALIIQVLLIKWSSCKKKLSLSTL 230  
DB 160 ANELRDSRLANDLRSGATIRIGIYGAGICAGLALIFGALIFKWSHSEKIQNLSTL 239  
QY 231 ITLANLPFGGLANAGAVIRSEENIYTIENVEYVENSNEYCYVNS-QQPS 281  
DB 240 ISLANLPBGLANAVABGIRSEENIYTIENVEYVEEENEYCYVNSQQPS 291

Search completed: November 22, 2003, 05:40:20  
Job time : 58.0412 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 22, 2003, 05:38:42 ; Search time 18.3471 Seconds  
(without alignments)  
648.024 Million cell updates/sec

Title: US-10-004-633-10

Perfect score: 1468

Sequence: 1 MFSGLTINCVLLTLQLLAR.....VEVENSNEYCYVNSQPS 281

Scoring table: BLOSUM62

Gapop 10.0 , Gapect 0.5

Searched: 328717 seqs, 4231058 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A COMB .pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B COMB .pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A COMB .pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B COMB .pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PCTUS COMB .pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/backfile1.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1468	100.0	281	1 US-08-487-748A-9	Sequence 9, Appl1
2	1468	100.0	281	1 US-08-487-748A-10	Sequence 10, Appl1
3	1468	100.0	281	3 US-08-480-070C-10	Sequence 10, Appl1
4	1468	100.0	281	3 US-08-829-525-10	Sequence 10, Appl1
5	1468	100.0	281	3 US-08-609-583A-10	Sequence 10, Appl1
6	1468	100.0	281	3 US-08-937-389-10	Sequence 10, Appl1
7	1468	100.0	281	4 US-09-310-367-10	Sequence 10, Appl1
8	1468	100.0	281	4 US-09-032-337-10	Sequence 10, Appl1
9	1468	100.0	281	4 US-09-464-231-10	Sequence 10, Appl1
10	883	60.1	301	3 US-08-829-525-24	Sequence 24, Appl1
11	883	60.1	301	3 US-08-937-389-24	Sequence 24, Appl1
12	883	60.1	301	3 US-09-310-367-24	Sequence 24, Appl1
13	883	60.1	301	4 US-09-032-337-24	Sequence 24, Appl1
14	883	60.1	301	4 US-09-464-231-24	Sequence 24, Appl1
15	883	60.1	301	4 US-09-464-231-24	Sequence 24, Appl1
16	267.5	18.2	451	1 US-08-287-001A-2	Sequence 2, Appl1
17	267.5	18.2	451	5 PCT-US95-09941-2	Sequence 2, Appl1
18	157	10.7	319	1 US-08-597-495B-22	Sequence 22, Appl1
19	157	10.7	319	3 US-09-068-051A-22	Sequence 22, Appl1
20	157	10.7	319	4 US-09-336-536-67	Sequence 6, Appl1
21	157	10.7	319	4 US-09-254-465A-6	Sequence 6, Appl1
22	151	10.3	273	4 US-09-254-465A-26	Sequence 26, Appl1
23	150	10.2	270	4 US-09-254-465A-24	Sequence 24, Appl1
24	149.5	10.2	365	4 US-08-928-383B-2	Sequence 2, Appl1
25	147.5	10.0	352	4 US-09-996-243-505	Sequence 505, App
26	147.5	10.0	365	2 US-08-979-424-3	Sequence 3, Appl1
27	147.5	10.0	365	3 US-09-272-496-2	Sequence 2, Appl1

28	143	9.7	318	3 US-09-068-051A-32	Sequence 32, Appl1
29	138	9.4	320	3 US-08-205-697A-2	Sequence 2, Appl1
30	138	9.4	320	3 US-08-702-525-2	Sequence 2, Appl1
31	138	9.4	320	5 PCT-US95-02576-2	Sequence 2, Appl1
32	135.5	9.2	466	4 US-09-604-107A-8	Sequence 8, Appl1
33	131	8.9	338	3 US-09-189-035-6	Sequence 6, Appl1
34	131	8.9	338	3 US-09-382-086-6	Sequence 6, Appl1
35	131	8.9	338	4 US-08-999-689A-5	Sequence 5, Appl1
36	131	8.9	503	4 US-08-999-689A-6	Sequence 6, Appl1
37	128.5	8.8	431	3 US-09-038-832-2	Sequence 2, Appl1
38	128.5	8.8	431	3 US-09-038-832-4	Sequence 4, Appl1
39	126.5	8.6	365	3 US-08-928-383B-23	Sequence 23, Appl1
40	126.5	8.6	365	3 US-08-928-383B-24	Sequence 24, Appl1
41	126.5	8.6	365	3 US-08-928-383B-26	Sequence 26, Appl1
42	124	8.4	306	2 US-08-147-172-4	Sequence 4, Appl1
43	124	8.4	306	2 US-08-456-104-8	Sequence 8, Appl1
44	124	8.4	306	2 US-08-101-624-25	Sequence 25, Appl1
45	124	8.4	306	3 US-08-153-262-4	Sequence 4, Appl1

## ALIGNMENTS

RESULT 1  
US-08-487-748A-9  
; Sequence 9, Application US/08487748A  
; Patent No. 5721351  
; GENERAL INFORMATION:  
; APPLICANT: Levanson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,748A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cornuzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-487-748A-9

Query Match 100.0%; Score 1468; DB 1; Length 281;  
Best Local Similarity 100.0%; Pred. No. 1.8e-150;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 MFSGLTINCVLLTLQLLARSLDGYKVEYKXNAVLPSCSTYLPSSGTLVPMCKGKGCPCW 60  
Db 1 MFSGLTINCVLLTLQLLARSLDGYKVEYKXNAVLPSCSTYLPSSGTLVPMCKGKGCPCW 60  
Cy 61 SGTNELLRTDERNNVTYOKSSRYOLKGLDKGVSLIIKKVTLDDHGTCCRIQPGIANN 120

Db 61 SQTNELRTDERNTYQKSSRYOLKGDINKGDSVLIINVTLLDHDGTYCCRIOPGLMN 120  
Qy 121 DKLELKLDIKAAKVTPAQTAGDSTTASPRITLTERNGSETOTLVTLHNNNGTKISTWA 180  
Db 121 DKLELKLDIKAAKVTPAQTAGDSTTASPRITLTERNGSETOTLVTLHNNNGTKISTWA 180  
Qy 181 DEIKSGEITRTAHIGVGSAGLTALIIIGVLIKWSCKKKLSLTLTIANLPPGG 240  
Db 181 DEIKSGEITRTAHIGVGSAGLTALIIIGVLIKWSCKKKLSLTLTIANLPPGG 240  
Qy 241 LANAGAVIRSEENITYTEENVEVENSNEYCYVNSQOPS 281  
Db 241 LANAGAVIRSEENITYTEENVEVENSNEYCYVNSQOPS 281

RESULT 2  
US-08-487-748A-10  
; Sequence 10, Application US/08487748A  
; Patent No. 5721351  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,748A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-487-748A-10

Query Match 100.0%; Score 1468; DB 1; Length 281;  
Best Local Similarity 100.0%; Pred. No. 1.8e-150;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 181 DEIKSGEITRTAHIGVGSAGLTALIIIGVLIKWSCKKKLSLTLTIANLPPGG 240  
Db 181 DEIKSGEITRTAHIGVGSAGLTALIIIGVLIKWSCKKKLSLTLTIANLPPGG 240  
Qy 241 LANAGAVIRSEENITYTEENVEVENSNEYCYVNSQOPS 281  
Db 241 LANAGAVIRSEENITYTEENVEVENSNEYCYVNSQOPS 281

RESULT 3  
US-08-480-070C-10  
; Sequence 10, Application US/08480070C  
; Patent No. 6066498  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/480,070C  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-024  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-480-070C-10

Query Match 100.0%; Score 1468; DB 3; Length 281;  
Best Local Similarity 100.0%; Pred. No. 1.8e-150;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
RESULT 4
US-08-829-525-10
; Sequence 10, Application US/08829525
; Patent No. 6084083
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/829,525
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,583
; FILING DATE: 01-MAR-1996
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-081
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-829-525-10

Query Match      100.0%; Score 1468; DB 3; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.8e-150;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSGLTINCVLLLLQLLILARSLEDEGYKVEGKNAVLPFCSYTLPTSGTLVPMCMGKGFPCPW 60
DB 1 MFSGLTINCVLLLLQLLILARSLEDEGYKVEGKNAVLPFCSYTLPTSGTLVPMCMGKGFPCPW 60
QY 61 SOCCTNELLRDERNNVYQKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIOPGLMN 120
DB 61 SOCCTNELLRDERNNVYQKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIOPGLMN 120
QY 121 DKKLELKLDIKAAKVTPTAOTAHGDSSTASPRITLTTERNGSETOTLVTLHNNNGTKISTWA 180
DB 121 DKKLELKLDIKAAKVTPTAOTAHGDSSTASPRITLTTERNGSETOTLVTLHNNNGTKISTWA 180
QY 181 DEIKDSEGTIRTAIHIGVGSAGITLALIIIGVLLIKWYSCKKKLSLSLITLANLPBG 240
DB 181 DEIKDSEGTIRTAIHIGVGSAGITLALIIIGVLLIKWYSCKKKLSLSLITLANLPBG 240
QY 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281
DB 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281
```

```
RESULT 5
US-08-609-583A-10
; Sequence 10, Application US/08609583A
; Patent No. 6204371
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,583A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-609-583A-10

Query Match      100.0%; Score 1468; DB 3; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.8e-150;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSGLTINCVLLLLQLLILARSLEDEGYKVEGKNAVLPFCSYTLPTSGTLVPMCMGKGFPCPW 60
DB 1 MFSGLTINCVLLLLQLLILARSLEDEGYKVEGKNAVLPFCSYTLPTSGTLVPMCMGKGFPCPW 60
QY 61 SOCCTNELLRDERNNVYQKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIOPGLMN 120
DB 61 SOCCTNELLRDERNNVYQKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIOPGLMN 120
QY 121 DKKLELKLDIKAAKVTPTAOTAHGDSSTASPRITLTTERNGSETOTLVTLHNNNGTKISTWA 180
DB 121 DKKLELKLDIKAAKVTPTAOTAHGDSSTASPRITLTTERNGSETOTLVTLHNNNGTKISTWA 180
QY 181 DEIKDSEGTIRTAIHIGVGSAGITLALIIIGVLLIKWYSCKKKLSLSLITLANLPBG 240
DB 181 DEIKDSEGTIRTAIHIGVGSAGITLALIIIGVLLIKWYSCKKKLSLSLITLANLPBG 240
QY 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281
DB 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281
```

## RESULT 6

US-08-937-399-10  
; Sequence 10, Application US/08937399  
; Patent No. 6288218  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/937,399  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-048  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-937-399-10

Query Match 100.0%; Score 1468; DB 3; Length 281;

Best Local Similarity 100.0%; Pred. No. 1.8e-150;

Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGSLTNCVLLIQLLLARSLDEGYKVEGKNAVLPSCSYTLPTSGTLVPMCMGKGFPM 60  
DB 1 MSGSLTNCVLLIQLLLARSLDEGYKVEGKNAVLPSCSYTLPTSGTLVPMCMGKGFPM 60  
QY 61 SQTNELLTDERNVTYQKSSRYQLKGDINKGDVSLIKNVTLDDHGTCCRIQPPGLMN 120  
DB 61 SQTNELLTDERNVTYQKSSRYQLKGDINKGDVSLIKNVTLDDHGTCCRIQPPGLMN 120  
QY 121 DKKLEKLDIAKAATPAQTAGDSTTASPRTLTTERNGSETOTLVTLHNNNGTKISTWA 180  
DB 121 DKKLEKLDIAKAATPAQTAGDSTTASPRTLTTERNGSETOTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDSGETIRTAHIGVGSAGLTALIIIGVLILKMWSCCKKKLSLITLANLPFG 240  
DB 181 DEIKDSGETIRTAHIGVGSAGLTALIIIGVLILKMWSCCKKKLSLITLANLPFG 240  
QY 241 LANAGAVRIRSEENIYTIENNVYEVENSNYYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENIYTIENNVYEVENSNYYCYVNSQOPS 281

## RESULT 7

US-09-310-367-10  
; Sequence 10, Application US/09310367  
; Patent No. 641417  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/310,367  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/829,525  
; FILING DATE: 28-MAR-1997  
; APPLICATION NUMBER: US 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-081  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-09-310-367-10

Query Match 100.0%; Score 1468; DB 4; Length 281;

Best Local Similarity 100.0%; Pred. No. 1.8e-150;

Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGSLTNCVLLIQLLLARSLDEGYKVEGKNAVLPSCSYTLPTSGTLVPMCMGKGFPM 60  
DB 1 MSGSLTNCVLLIQLLLARSLDEGYKVEGKNAVLPSCSYTLPTSGTLVPMCMGKGFPM 60  
QY 61 SQTNELLTDERNVTYQKSSRYQLKGDINKGDVSLIKNVTLDDHGTCCRIQPPGLMN 120  
DB 61 SQTNELLTDERNVTYQKSSRYQLKGDINKGDVSLIKNVTLDDHGTCCRIQPPGLMN 120  
QY 121 DKKLEKLDIAKAATPAQTAGDSTTASPRTLTTERNGSETOTLVTLHNNNGTKISTWA 180  
DB 121 DKKLEKLDIAKAATPAQTAGDSTTASPRTLTTERNGSETOTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDSGETIRTAHIGVGSAGLTALIIIGVLILKMWSCCKKKLSLITLANLPFG 240  
DB 181 DEIKDSGETIRTAHIGVGSAGLTALIIIGVLILKMWSCCKKKLSLITLANLPFG 240  
QY 241 LANAGAVRIRSEENIYTIENNVYEVENSNYYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENIYTIENNVYEVENSNYYCYVNSQOPS 281

## RESULT 8

US-09-032-337-10  
; Sequence 10, Application US/09032337  
; Patent No. 6455685  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/032.337  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cortuzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-016  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-09-032-337-10

Query Match 100.0%; Score 1468; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 1.8e-150;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFSGLTINCVLLIQLLLARSLEDEGYVEVGKNAVLPSCSYTLPTSGTLVPMCMGKGFCPW 60  
Db 1 MFSGLTINCVLLIQLLLARSLEDEGYVEVGKNAVLPSCSYTLPTSGTLVPMCMGKGFCPW 60  
QY 61 SÖCTNELLRDERNNVYÖKSSRYÖLKDNLKGDVSLIIKNVTLDDHGTCCRIÖPGLMN 120  
Db 61 SÖCTNELLRDERNNVYÖKSSRYÖLKDNLKGDVSLIIKNVTLDDHGTCCRIÖPGLMN 120  
QY 121 DKKLELKDIDKAAKVPAPQTAHGDSTTASPRTLTTERNGSETÖTLVTLHNNNGTKISTWA 180  
Db 121 DKKLELKDIDKAAKVPAPQTAHGDSTTASPRTLTTERNGSETÖTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDGETIRTAIHIGVGSAGLTALIIIGVLLIKWYSCKKKLSLSITTLANLPGG 240  
Db 181 DEIKDGETIRTAIHIGVGSAGLTALIIIGVLLIKWYSCKKKLSLSITTLANLPGG 240  
QY 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281  
Db 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281

## RESULT 9

US-09-464-231-10  
; Sequence 10, Application US/09464231  
; Patent No. 6562343  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/464,231  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cortuzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-048  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-09-464-231-10

Query Match 100.0%; Score 1468; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 1.8e-150;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFSGLTINCVLLIQLLLARSLEDEGYVEVGKNAVLPSCSYTLPTSGTLVPMCMGKGFCPW 60  
Db 1 MFSGLTINCVLLIQLLLARSLEDEGYVEVGKNAVLPSCSYTLPTSGTLVPMCMGKGFCPW 60  
QY 61 SÖCTNELLRDERNNVYÖKSSRYÖLKDNLKGDVSLIIKNVTLDDHGTCCRIÖPGLMN 120  
Db 61 SÖCTNELLRDERNNVYÖKSSRYÖLKDNLKGDVSLIIKNVTLDDHGTCCRIÖPGLMN 120  
QY 121 DKKLELKDIDKAAKVPAPQTAHGDSTTASPRTLTTERNGSETÖTLVTLHNNNGTKISTWA 180  
Db 121 DKKLELKDIDKAAKVPAPQTAHGDSTTASPRTLTTERNGSETÖTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDGETIRTAIHIGVGSAGLTALIIIGVLLIKWYSCKKKLSLSITTLANLPGG 240  
Db 181 DEIKDGETIRTAIHIGVGSAGLTALIIIGVLLIKWYSCKKKLSLSITTLANLPGG 240  
QY 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281  
Db 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281

RESULT 10  
US-08-829-525-24  
Sequence 24, Application US/08829525  
Patent No. 6084083

## GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:

ADDRESS: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA

ZIP: 10036/2711

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/829,525  
FILING DATE: 28-MAR-1997

CLASSIFICATION: 536  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996

APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7853-081  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:

LENGTH: 301 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein  
FRAGMENT TYPE: internal

US-08-829-525-24

Query Match 60.1%; Score 883; DB 3; Length 301;  
Best Local Similarity 63.7%; Pred. No. 3.9e-87;

Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MMSGTLNVCVLLDQLLLARSLSDGYKVEGKNAIPLCSYTLPTSGTIVPMCKGFCPCW 60  
DB 1 MESHLPFDCVLLLLLTTRSSSEVEYRAVQONAYLPCEYTPAPAGNLVPCWKGACPCV 60  
QY 61 SOCTNELRTDERNTYOKSSRYOLKGDINKGVSLLIKNVTLDHGTCCRIOPGLMN 120  
DB 61 PECGNVLTIRDERDNYM-TSRWMLNGDPRKGDVSLTIENVTLADSGIYCCRIQIPGINN 119  
QY 121 DKKLBIKLDIKAKVTPAQTAGDSTTASPRTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
DB 120 DEKFNKLKVIKAKVTPAQTAGDSTTASPRTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ADE-----IKSGETITRAIHIGVVSAGLTALIIIGVLLKWTSCCKKUSLSL 230  
DB 180 ANELRDSRLANDLRDSDGATIRIGIYIGAGICAGLALALIFGALIFKWSHSEKIONLSL 239  
QY 231 ITLANLPPGLNAGAVIRISEENITYIEENVYEVENSNEYCYVNS-QQPS 281

Db 240 ISLANLPPSGLNAGAVIRISEENITYIEENVYEVENSNEYCYVNS-QQPS 291

RESULT 11  
US-08-609-583A-24  
Sequence 24, Application US/08609583A  
Patent No. 6204371

## GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:

ADDRESS: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA

ZIP: 10036/2711

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/609,583A  
FILING DATE: 01-MAR-1996

CLASSIFICATION: 536  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:

LENGTH: 301 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein  
FRAGMENT TYPE: internal

US-08-609-583A-24

Query Match 60.1%; Score 883; DB 3; Length 301;  
Best Local Similarity 63.7%; Pred. No. 3.9e-87;

Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MMSGTLNVCVLLDQLLLARSLSDGYKVEGKNAIPLCSYTLPTSGTIVPMCKGFCPCW 60  
DB 1 MESHLPFDCVLLLLLTTRSSSEVEYRAVQONAYLPCEYTPAPAGNLVPCWKGACPCV 60  
QY 61 SOCTNELRTDERNTYOKSSRYOLKGDINKGVSLLIKNVTLDHGTCCRIOPGLMN 120  
DB 61 PECGNVLTIRDERDNYM-TSRWMLNGDPRKGDVSLTIENVTLADSGIYCCRIQIPGINN 119  
QY 121 DKKLBIKLDIKAKVTPAQTAGDSTTASPRTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
DB 120 DEKFNKLKVIKAKVTPAQTAGDSTTASPRTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ADE-----IKSGETITRAIHIGVVSAGLTALIIIGVLLKWTSCCKKUSLSL 230  
DB 180 ANELRDSRLANDLRDSDGATIRIGIYIGAGICAGLALALIFGALIFKWSHSEKIONLSL 239  
QY 231 ITLANLPPGLNAGAVIRISEENITYIEENVYEVENSNEYCYVNS-QQPS 281

Db 240 ISLANLPFSGLANAVALGIRSEENITYIIEENVYEVERPEPNEYCYVSSRQOPS 291

RESULT 12  
US-08-937-399-24

; Sequence 24, Application US/08937399  
; Patent No. 6288218  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/937,399  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-048  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 301 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; US-08-937-399-24

Query Match 60.1%; Score 883; DB 3; Length 301;  
Best Local Similarity 63.7%; Pred. No. 3.9e-87;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

Qy 1 MFSGLTINCVALLIQILLASLEDGYKVEGKNAVLPSCSYTLPTSGTLVPMCKGKGFPM 60  
Db 1 MFSHLPDVCVLLILLTLTSSEVEYRAEVQONAYLPCFYTPAAGMLVPCWKGKACPV 60  
Qy 61 SOCNELLRTERDEVNTYQKSSRYOLKGDINKGVSLIKKVTLLDDHGTVCRIQIPGIAM 120  
Db 61 PEGGNVLRTERDEVNTW-TSRVWLNDFPKGVDVSLTEENVTLADSGIYCCRIQIPGIAM 119  
Qy 121 DKLELKLIDIKAAKVTPAQTAAHGDSTTASPRTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
Db 120 DEKNLKLVIKPAKVTAPTLQROFTAAFPRLTTRHGGAPEQTGLSLPINDITQISTL 179  
Qy 180 ADE-----IDSGETIRTAIHIGVAGAGLTALIIIVLILKMYSCCKKULSSLSL 230  
Db 180 ANELRDSRLANDLRDGSATIRIGIYIGAGICAGLALIFGALLFKWYSHSEKIKONLSL 239

Qy 231 ITLANLPFGLANAGAVRIRSEENITYIIEENVYEVENSNEYCYVNS-QOPS 281  
Db 240 ISLANLPFSGLANAVALGIRSEENITYIIEENVYEVERPEPNEYCYVSSRQOPS 291

RESULT 13  
US-09-310-367-24

; Sequence 24, Application US/09310367  
; Patent No. 6414117  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/310,367  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/829,525  
; FILING DATE: 28-MAR-1997  
; APPLICATION NUMBER: US 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-081  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 301 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; US-09-310-367-24

Query Match 60.1%; Score 883; DB 4; Length 301;  
Best Local Similarity 63.7%; Pred. No. 3.9e-87;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

Qy 1 MFSGLTINCVALLIQILLASLEDGYKVEGKNAVLPSCSYTLPTSGTLVPMCKGKGFPM 60  
Db 1 MFSHLPDVCVLLILLTLTSSEVEYRAEVQONAYLPCFYTPAAGMLVPCWKGKACPV 60  
Qy 61 SOCNELLRTERDEVNTYQKSSRYOLKGDINKGVSLIKKVTLLDDHGTVCRIQIPGIAM 120  
Db 61 PEGGNVLRTERDEVNTW-TSRVWLNDFPKGVDVSLTEENVTLADSGIYCCRIQIPGIAM 119  
Qy 121 DKLELKLIDIKAAKVTPAQTAAHGDSTTASPRTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
Db 120 DEKNLKLVIKPAKVTAPTLQROFTAAFPRLTTRHGGAPEQTGLSLPINDITQISTL 179

[illegible]

RESULT 14

US-09-032-337-24  
Sequence 24, Application US/09032337  
Patent No. 6455685  
GENERAL INFORMATION:  
APPLICANT: levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/032,337  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-032-337-24

Query Match	60.1%;	Score 883;	DB 4;	Length 301;
Best Local Similarity	63.7%;	Pred. No. 3.9e-87;		
Matches 186;	Conservative 26;	Mismatches 68;	Indels 12;	Gaps 4
QY	1	MFSGLVLTNCTLLTLLQLLARSLDEGKYKVEVGNATLPFCSTYLLPFSGLTVPMCKGKGC	60	
Db	1	MFSLHPDPCLLTLLTLLTTRSSSEVERAEVGNATLPFCYTPAAPGNLVPMCKGKGC	60	
QY	61	SOCTNELLTPDENATVYQKSSRYQLKGLNKGDVSLIKANTLDDHGVCRCRIOPSLMN	120	
Db	61	FECGNVVLRTDERDVYVM-TSRVYMGJGFRKGDVSLTIENVTLLDSGVCRCRIOPSLMN	119	
QY	121	DKLLELKLDIKAKVTPAOTAGDSVTASPRLLTTERNG-SETQTLVTLHNHNTKISTW	179	

[illegible]

RESULT 15  
ME-09-164

Sequence 24, Application US/09464231  
Patent No. 6562343  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/464,231  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
US-09-464-231-24

	Query Match	60.1%	Score 883;	DB 4;	Length 301;	
	Best Local Similarity	63.7%	Pred. No.	3,9e-87;		
	Matches	166;	Conservative	26;	Mismatches 68;	Indels 12; Gaps 4
QY	1	MSGGLTNCVLLLTQLLARSLEDGYKRYENKNAYLPDSCSYTLPTPSGTLYPMCKGRGCPW	60			
		:           :           :           :           :           :				
		:           :           :           :           :				
		:           :           :           :           :				
Dd	1	MSHHPDCVCLLLELLLRSEVEFRREVQNNAIPFTTPAAGLVPEVCWKAGACPV	60			
		:           :           :           :           :				
		:           :           :           :           :				
		:           :           :           :           :				
QY	61	SQCETNELLTRERNATYQSSRRYLKGDLNGADVSLIIKNVTLDHGTYCCRIQFPLGN	120			
		:           :           :           :           :				
		:           :           :           :           :				
		:           :           :           :           :				
	61	FECGNVNLTEEDRDNYVM-TSRMYLNMGFFRGADSVLIETNVLADSGIYYCCRIOEIPINN	119			
		:           :           :           :           :				
		:           :           :           :           :				
		:           :           :           :           :				

QY 121 DKKELEKLDIKAQVTPAQTAHGDSTASPRILTERNG-SETQTLVTLHNNNGTKISTW 179  
 Db 120 DEKFNLKLVTKPAKVTPAPTLQORDFTAPPRMLTTRGHGPAETQTLGSLPDINLTQISTL 179  
 QY 180 ADE-----IKDSEITRTAIHIGVGSAGITLALIIGVLILKWSCKKKLSL 230  
 Db 180 ANELRDSRLANDRDSGATTIRIGIYGAGICAGLALIFGALIFKWSHSHKEKIQNL 239  
 QY 231 ITLANLPFGGLANAGAVRIRSEENITYIEENVYEVENSNEYCYVNS-OOPS 281  
 Db 240 ISIANLPFGGLANAVAGIRSEENITYIEENVYEVENPNBYCYVSSROOPS 291

Search completed: November 22, 2003, 05:45:19  
 Job time : 25.3471 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 05:43:31 ; Search time 42.0052 Seconds  
(without alignments)  
1221.260 Million cell updates/sec

Title: US-10-004-633-10

Perfect score: 1468  
Sequence: 1 MFSLTLNVCVLLQLLLAR.....VVEVENENYCYVNSOOPS 281

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10C\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1468	100.0	281	US-10-252-131-10	Sequence 10, Appl
2	1468	100.0	281	US-10-004-633-10	Sequence 10, Appl
3	1468	100.0	281	US-10-188-012-9	Sequence 9, Appl
4	1427	97.2	281	US-10-188-012-11	Sequence 11, Appl
5	883	60.1	301	US-10-252-131-24	Sequence 24, Appl
6	883	60.1	301	US-10-004-633-24	Sequence 24, Appl
7	883	60.1	301	US-10-188-012-29	Sequence 29, Appl
8	883	60.1	301	US-10-188-012-31	Sequence 31, Appl
9	520	35.4	185	US-09-529-063-34	Sequence 34, Appl
10	430.5	29.3	142	US-10-414-378-34	Sequence 34, Appl
11	430.5	29.3	142	US-09-966-546-22	Sequence 22, Appl
12	430.5	29.3	142	US-09-966-546-22	Sequence 22, Appl
13	430.5	29.3	142	US-09-965-212-22	Sequence 22, Appl
14	430.5	29.3	142	US-10-189-940-22	Sequence 22, Appl
15	430.5	29.3	142	US-10-189-940-146	Sequence 146, App

16	374	25.5	282	US-10-188-012-3	Sequence 3, Appl
17	367.5	25.0	125	US-10-189-940-145	Sequence 145, App
18	364.5	24.8	305	US-10-188-012-1	Sequence 1, Appl
19	328.5	22.4	359	US-10-188-012-17	Sequence 17, Appl
20	326.5	22.2	359	US-10-188-012-19	Sequence 19, Appl
21	326	22.2	305	US-10-188-012-5	Sequence 5, Appl
22	326	22.2	305	US-10-188-012-7	Sequence 7, Appl
23	324	22.1	364	US-10-188-012-25	Sequence 25, Appl
24	323.5	22.0	365	US-10-188-012-21	Sequence 21, Appl
25	321.5	21.9	359	US-10-188-012-23	Sequence 23, Appl
26	316	21.5	364	US-10-188-012-27	Sequence 27, Appl
27	276.5	18.8	345	US-10-188-012-13	Sequence 13, Appl
28	275	18.7	345	US-10-188-012-15	Sequence 15, Appl
29	274.5	18.7	183	US-09-739-907-65	Sequence 65, Appl
30	274.5	18.7	378	US-10-188-012-33	Sequence 33, Appl
31	274.5	18.7	378	US-10-188-012-35	Sequence 35, Appl
32	274.5	18.7	379	US-10-188-012-37	Sequence 37, Appl
33	157	10.7	319	US-09-813-153-138	Sequence 138, App
34	157	10.7	319	US-09-813-153-138	Sequence 138, App
35	157	10.7	319	US-09-813-153-138	Sequence 138, App
36	154.5	10.5	323	US-10-265-542-6	Sequence 6, Appl
37	154.5	10.5	323	US-10-265-542-6	Sequence 6, Appl
38	151	10.3	273	US-09-971-798-27	Sequence 27, Appl
39	151	10.3	273	US-09-953-499-26	Sequence 26, Appl
40	150	10.2	268	US-10-265-542-24	Sequence 24, Appl
41	150	10.2	270	US-10-265-542-24	Sequence 24, Appl
42	148	10.1	505	US-09-953-499-24	Sequence 24, Appl
43	147.5	10.0	352	US-10-114-153-12	Sequence 12, Appl
44	147.5	10.0	352	US-09-989-722-505	Sequence 505, App
45	147.5	10.0	352	US-09-989-723-505	Sequence 505, App
			9	US-09-989-279-505	Sequence 505, App

## ALIGNMENTS

RESULT 1  
US-10-252-131-10  
Sequence 10, Application US/10252131  
Publication No. US20030158399A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/252,131  
FILING DATE: 20-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-10-252-131-10

Query Match 100.0%; Score 1468; DB 12; Length 281;  
Best Local Similarity 100.0%; Pred. No. 5.2e-127;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGSLTNCVLLLLQLLARSLSDGYKVEGKNAVYPCSYTLPSTGLVPMCGKGFCEW 60  
DB 1 MSGSLTNCVLLLLQLLARSLSDGYKVEGKNAVYPCSYTLPSTGLVPMCGKGFCEW 60  
QY 61 SOCTNELRTDERNTYQKSSRYQLKGNKDVSLIKNVTLDHGYCCRIQPPGLMN 120  
DB 61 SOCTNELRTDERNTYQKSSRYQLKGNKDVSLIKNVTLDHGYCCRIQPPGLMN 120  
QY 121 DKLELKDIDIAKAVTPAQTAGDSTTASPRITLTERNGSETQTLVTLHNNNGTKISTWA 180  
DB 121 DKLELKDIDIAKAVTPAQTAGDSTTASPRITLTERNGSETQTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDSGETIRTAIHIGVGSAGLTALIIIGVLIKWSCKKKLSLSLITLANLPPGG 240  
DB 181 DEIKDSGETIRTAIHIGVGSAGLTALIIIGVLIKWSCKKKLSLSLITLANLPPGG 240  
QY 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281

RESULT 2  
US-10-004-633-10

Sequence 10, Application US/10004633  
Publication No. US20030069196A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
APPLICANT: Lloyd, Clare M.  
APPLICANT: McCarthy, Sean A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
FILE REFERENCE: 7853-125  
CURRENT APPLICATION NUMBER: US/10/004,633  
PRIOR FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,563  
PRIOR FILING DATE: EARLIER FILING DATE: 1996-03-01  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/487,748  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/398,633  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-03-05  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 10  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-004-633-10

Query Match 100.0%; Score 1468; DB 15; Length 281;  
Best Local Similarity 100.0%; Pred. No. 5.2e-127;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGSLTNCVLLLLQLLARSLSDGYKVEGKNAVYPCSYTLPSTGLVPMCGKGFCEW 60  
DB 1 MSGSLTNCVLLLLQLLARSLSDGYKVEGKNAVYPCSYTLPSTGLVPMCGKGFCEW 60

QY 61 SOCTNELRTDERNTYQKSSRYQLKGNKDVSLIKNVTLDHGYCCRIQPPGLMN 120  
DB 61 SOCTNELRTDERNTYQKSSRYQLKGNKDVSLIKNVTLDHGYCCRIQPPGLMN 120  
QY 121 DKLELKDIDIAKAVTPAQTAGDSTTASPRITLTERNGSETQTLVTLHNNNGTKISTWA 180  
DB 121 DKLELKDIDIAKAVTPAQTAGDSTTASPRITLTERNGSETQTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDSGETIRTAIHIGVGSAGLTALIIIGVLIKWSCKKKLSLSLITLANLPPGG 240  
DB 181 DEIKDSGETIRTAIHIGVGSAGLTALIIIGVLIKWSCKKKLSLSLITLANLPPGG 240  
QY 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281

RESULT 3  
US-10-188-012-9

Sequence 9, Application US/10188012  
Publication No. US20030124114A1  
GENERAL INFORMATION:  
APPLICANT: McIntire, Jennifer Jones  
APPLICANT: Umetsu, Dale T.  
APPLICANT: Dekruyff, Rosemarie  
APPLICANT: Kuchroo, Vijay  
APPLICANT: Freeman, Gordon J.  
TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
FILE REFERENCE: STAN-235  
CURRENT APPLICATION NUMBER: US/10/188,012  
PRIOR FILING DATE: 2002-07-01  
PRIOR APPLICATION NUMBER: 60/302,344  
PRIOR FILING DATE: 2001-06-29  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(281)  
OTHER INFORMATION: TIM-3 BALB/c allele  
US-10-188-012-9

Query Match 100.0%; Score 1468; DB 15; Length 281;  
Best Local Similarity 100.0%; Pred. No. 5.2e-127;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGSLTNCVLLLLQLLARSLSDGYKVEGKNAVYPCSYTLPSTGLVPMCGKGFCEW 60  
DB 1 MSGSLTNCVLLLLQLLARSLSDGYKVEGKNAVYPCSYTLPSTGLVPMCGKGFCEW 60  
QY 61 SOCTNELRTDERNTYQKSSRYQLKGNKDVSLIKNVTLDHGYCCRIQPPGLMN 120  
DB 61 SOCTNELRTDERNTYQKSSRYQLKGNKDVSLIKNVTLDHGYCCRIQPPGLMN 120  
QY 121 DKLELKDIDIAKAVTPAQTAGDSTTASPRITLTERNGSETQTLVTLHNNNGTKISTWA 180  
DB 121 DKLELKDIDIAKAVTPAQTAGDSTTASPRITLTERNGSETQTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDSGETIRTAIHIGVGSAGLTALIIIGVLIKWSCKKKLSLSLITLANLPPGG 240  
DB 181 DEIKDSGETIRTAIHIGVGSAGLTALIIIGVLIKWSCKKKLSLSLITLANLPPGG 240  
QY 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281

RESULT 4  
US-10-188-012-11

Sequence 11, Application US/10188012  
Publication No. US20030124114A1  
GENERAL INFORMATION:  
APPLICANT: Mcintire, Jennifer Jones  
APPLICANT: Umetsu, Dale T.  
APPLICANT: Dekruyf, Rosemarie  
APPLICANT: Kuchroo, Vijay  
APPLICANT: Freeman, Gordon J.  
TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
TITLE OF INVENTION: Use Thereof  
FILE REFERENCE: STAN-235  
CURRENT APPLICATION NUMBER: US/10/188,012  
CURRENT FILING DATE: 2002-07-01  
PRIOR APPLICATION NUMBER: 60/302,344  
PRIOR FILING DATE: 2001-06-29  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for windows Version 4.0  
SEQ ID NO 11  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(281)  
OTHER INFORMATION: TIM-3, C.D2 ES-HBA and DBA/2J allele  
US-10-188-012-11

Query Match 97.2%; Score 1427; DB 15; Length 281;  
Best Local Similarity 97.5%; Pred. No. 3,1e-123;  
Matches 274; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 MFSGLTNCVLLIQLLLARSLDGYKVEGKNAVLPCTSYTLPTSGTLVPMCKGKFCPW 60  
DB 1 MFSGLTNCVLLIQLLLARSLDGYKVEGKNAVLPCTSYTLPTSGTLVPMCKGKFCPW 60  
QY 61 SQTNELRLTDERNVYQKSRVYOLKGDNLKGDVSLIKNVTLDHGTCCRIQPGIMN 120  
DB 61 SQTNELRLTDERNVYQKSRVYOLKGDNLKGDVSLIKNVTLDHGTCCRIQPGIMN 120  
QY 121 DKLELKLIDIKAAVTPAQTADHGSTASPTLTTERNGSETOTLVTLHNNNGTKISTWA 180  
DB 121 DKLELKLIDIKAAVTPAQTADHGSTASPTLTTERNGSETOTLVTLHNNNGTKISTWA 180  
QY 181 DEIKSGETIRTAHIGVGSAGLTLLIIGVLLKMYSCCKKKLSLSTLTLANLPBG 240  
DB 181 DEIKSGETIRTAHIGVGSAGLTLLIIGVLLKMYSCCKKKLSLSTLTLANLPBG 240  
QY 241 LANAGAVRIRSEENIYTIENNVYEVENSNEYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENIYTIENNVYEVENSNEYCYVNSQOPS 281

RESULT 5  
US-10-252-131-24  
Sequence 24, Application US/10252131  
Publication No. US20030158399A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/252,131  
FILING DATE: 20-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 24:  
US-10-252-131-24

Query Match 60.1%; Score 883; DB 12; Length 301;  
Best Local Similarity 63.7%; Pred. No. 4,4e-73;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSGLTNCVLLIQLLLARSLDGYKVEGKNAVLPCTSYTLPTSGTLVPMCKGKFCPW 60  
DB 1 MFSGLTNCVLLIQLLLARSLDGYKVEGKNAVLPCTSYTLPTSGTLVPMCKGKFCPW 60  
QY 61 SQTNELRLTDERNVYQKSRVYOLKGDNLKGDVSLIKNVTLDHGTCCRIQPGIMN 120  
DB 61 SQTNELRLTDERNVYQKSRVYOLKGDNLKGDVSLIKNVTLDHGTCCRIQPGIMN 120  
QY 121 DKLELKLIDIKAAVTPAQTADHGSTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
DB 121 DKLELKLIDIKAAVTPAQTADHGSTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
QY 180 ADE-----IKDSEGTIRTAHIGVGSAGLTLLIIGVLLKMYSCCKKKLSLST 230  
DB 180 ADE-----IKDSEGTIRTAHIGVGSAGLTLLIIGVLLKMYSCCKKKLSLST 230  
QY 231 ITLANLPBGGLANAGAVRIRSEENIYTIENNVYEVENSNEYCYVNS-QOPS 281  
DB 231 ITLANLPBGGLANAGAVRIRSEENIYTIENNVYEVENSNEYCYVNS-QOPS 281  
QY 240 ISLANLPBGGLANAGAVRIRSEENIYTIENNVYEVENSNEYCYVNSRQOPS 291  
DB 240 ISLANLPBGGLANAGAVRIRSEENIYTIENNVYEVENSNEYCYVNSRQOPS 291

RESULT 6  
US-10-004-633-24  
Sequence 24, Application US/10004633  
Publication No. US20030069196A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
APPLICANT: Lloyd, Clare M.  
APPLICANT: McCarthy, Sean A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
DIAGNOSIS OF IMMUNE DISORDERS  
FILE REFERENCE: 7853-125  
CURRENT APPLICATION NUMBER: US/10/004,633  
CURRENT FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: 09/324,986  
PRIOR FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: 08/609,583  
PRIOR FILING DATE: 1996-03-01  
PRIOR APPLICATION NUMBER: 08/487,748



```

; Sequence 34, Application US/09529063
; Patent No. US20020102542A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, DAICHI
; APPLICANT: SHIBAYAMA, SHIRO
; APPLICANT: TADA, HIDEAKI
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF
; FILE REFERENCE: 058769
; CURRENT APPLICATION NUMBER: US/09/529,063
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: PCT/JP98/04514
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: JP 9-274674
; PRIOR FILING DATE: 1997-10-07
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-063-34

```

```

Query Match
Best Local Similarity 62.1%; Score 520; DB 10; Length 185;
Matches 110; Conservative 15; Mismatches 50; Indels 2; Gaps 2;

```

```

Qy 1 MFSGLTNCVLLQLLLARSLSDGYKVEGKNAVLPSCSYTLPTSGTLVPMCKGKCPW 60
Db 1 MFSHPDPCVLLLLLLLRSSSEVEYRAVGONAYLPFCFYTPAAPGNLVPVCKGKACP 60
Qy 61 SOCTNELRTDERNVYQKSSRYOLKGDVSLIKVNTLDDHGTCCRIOPGLMN 120
Db 61 PECGNVLRTERDENVYV-TSRVWLNQDFRKGDVSLITENTVLADSGIYCCRIQPIGMN 119
Qy 121 DKLELKLIDIKAKVTPAQTAHGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKI 176
Db 120 DEKNLKLVIKPAKVPAPLTQRDFTAFAFPRMLTTRGHGPAETQTLGSLPDINLTGI 176

```

```

RESULT 10
US-10-414-378-34
; Sequence 34, Application US/10414378
; Publication No. US20030165981A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, DAICHI
; APPLICANT: SHIBAYAMA, SHIRO
; APPLICANT: TADA, HIDEAKI
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF
; FILE REFERENCE: 058769
; CURRENT APPLICATION NUMBER: US/10/414,378
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/09/529,063
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: PCT/JP98/04514
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: JP 9-274674
; PRIOR FILING DATE: 1997-10-07
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-414-378-34

```

```

Query Match
Best Local Similarity 62.1%; Score 520; DB 12; Length 185;
Matches 110; Conservative 15; Mismatches 50; Indels 2; Gaps 2;
1 MFSGLTNCVLLQLLLARSLSDGYKVEGKNAVLPSCSYTLPTSGTLVPMCKGKCPW 60

```

```

Db 1 MFSHPDPCVLLLLLLLRSSSEVEYRAVGONAYLPFCFYTPAAPGNLVPVCKGKACP 60
Qy 61 SOCTNELRTDERNVYQKSSRYOLKGDVSLIKVNTLDDHGTCCRIOPGLMN 120
Db 61 PECGNVLRTERDENVYV-TSRVWLNQDFRKGDVSLITENTVLADSGIYCCRIQPIGMN 119
Qy 121 DKLELKLIDIKAKVTPAQTAHGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKI 176
Db 120 DEKNLKLVIKPAKVPAPLTQRDFTAFAFPRMLTTRGHGPAETQTLGSLPDINLTGI 176

```

```

RESULT 11
US-09-966-546-22
; Sequence 22, Application US/09966546
; Patent No. US20020168716A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20020168716A1el Human Proteins and Polynucleotides
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,546
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-546-22

```

```

Query Match
Best Local Similarity 63.2%; Score 430.5; DB 10; Length 142;
Matches 86; Conservative 12; Mismatches 37; Indels 1; Gaps 1;

```

```

Qy 1 MFSGLTNCVLLQLLLARSLSDGYKVEGKNAVLPSCSYTLPTSGTLVPMCKGKCPW 60
Db 1 MFSHPDPCVLLLLLLLRSSSEVEYRAVGONAYLPFCFYTPAAPGNLVPVCKGKACP 60
Qy 61 SOCTNELRTDERNVYQKSSRYOLKGDVSLIKVNTLDDHGTCCRIOPGLMN 120
Db 61 PECGNVLRTERDENVYV-TSRVWLNQDFRKGDVSLITENTVLADSGIYCCRIQPIGMN 119
Qy 121 DKLELKLIDIKAKVT 136
Db 120 DEKNLKLVIKPGEMT 135

```

```

RESULT 12
US-09-966-545-22
; Sequence 22, Application US/09966545
; Patent No. US20020172999A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20020172999A1el Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,545
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-545-22

```

```
Query Match      29.3%; Score 430.5; DB 10; Length 142;
Best Local Similarity 63.2%; Pred. No. 7.4e-32;
Matches      86; Conservative 12; Mismatches 37; Indels 1; Gaps 1;

QY      1 MFSGLTNCVLLLLQLLARSLEDGYKVEVGKNAVLPSCSYTLPTSGTLVPMCGKGFPCW 60
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      1 MFSHLPFDCVLLLLLLLRSSSEVEBRAVGONAYLPCTYTPAAGNLVPCMGKACPV 60
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY      61 SOCTNELRTDERNTYQKSSRYQLKGDINKGDSLIINKVTLDDHGTCCRIQPFGLMN 120
        :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      61 FECGNVVLRTDERDVYWM-TSRWYLNGBDFRKGDVSLTIENVTLADSGIYCCRIQIFGINN 119
        :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY      121 DKLELKLDIKAKYV 136
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      120 DEKFNKLVIKPEWT 135
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
US-09-965-212-22
; Sequence 22, Application US/09965212
; Publication No. US20030003462A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20030003462A1el Human Proteins and Polynucleotides Encoding
; TITLE OF INVENTION: Them
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/965,212
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: USSN 60/128,514
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 22
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-212-22

Query Match      29.3%; Score 430.5; DB 11; Length 142;
Best Local Similarity 63.2%; Pred. No. 7.4e-32;
Matches      86; Conservative 12; Mismatches 37; Indels 1; Gaps 1;

QY      1 MFSGLTNCVLLLLQLLARSLEDGYKVEVGKNAVLPSCSYTLPTSGTLVPMCGKGFPCW 60
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      1 MFSHLPFDCVLLLLLLLRSSSEVEBRAVGONAYLPCTYTPAAGNLVPCMGKACPV 60
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY      61 SOCTNELRTDERNTYQKSSRYQLKGDINKGDSLIINKVTLDDHGTCCRIQPFGLMN 120
        :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      61 FECGNVVLRTDERDVYWM-TSRWYLNGBDFRKGDVSLTIENVTLADSGIYCCRIQIFGINN 119
        :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY      121 DKLELKLDIKAKYV 136
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      120 DEKFNKLVIKPEWT 135
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
US-10-189-940-22
; Sequence 22, Application US/10189940
; Publication No. US20030129613A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard
; APPLICANT: Anderson, David
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh
```

```
APPLICANT: Casman, Stacie
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: No. US20030129613A1el Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: 15966-546 CIP
; CURRENT APPLICATION NUMBER: US/10/189,940
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/303,241
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/369,065
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/378,730
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 09/965,212
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,546
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/186,592
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 22
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-189-940-22

Query Match      29.3%; Score 430.5; DB 16; Length 142;
Best Local Similarity 63.2%; Pred. No. 7.4e-32;
Matches      86; Conservative 12; Mismatches 37; Indels 1; Gaps 1;

QY      1 MFSGLTNCVLLLLQLLARSLEDGYKVEVGKNAVLPSCSYTLPTSGTLVPMCGKGFPCW 60
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      1 MFSHLPFDCVLLLLLLLRSSSEVEBRAVGONAYLPCTYTPAAGNLVPCMGKACPV 60
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY      61 SOCTNELRTDERNTYQKSSRYQLKGDINKGDSLIINKVTLDDHGTCCRIQPFGLMN 120
        :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      61 FECGNVVLRTDERDVYWM-TSRWYLNGBDFRKGDVSLTIENVTLADSGIYCCRIQIFGINN 119
        :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY      121 DKLELKLDIKAKYV 136
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      120 DEKFNKLVIKPEWT 135
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 15
US-10-189-940-146
; Sequence 146, Application US/10189940
; Publication No. US20030129613A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard
; APPLICANT: Anderson, David
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh
; APPLICANT: Casman, Stacie
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: No. US20030129613A1el Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: 15966-546 CIP
; CURRENT APPLICATION NUMBER: US/10/189,940
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/303,241
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/369,065
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/378,730
```

; PRIOR FILING DATE: 2002-05-08  
 ; PRIOR APPLICATION NUMBER: 09/965,212  
 ; PRIOR FILING DATE: 2001-09-26  
 ; PRIOR APPLICATION NUMBER: 09/966,545  
 ; PRIOR FILING DATE: 2001-09-26  
 ; PRIOR APPLICATION NUMBER: 09/966,546  
 ; PRIOR FILING DATE: 2001-09-26  
 ; PRIOR APPLICATION NUMBER: 09/544,511  
 ; PRIOR FILING DATE: 2000-04-06  
 ; PRIOR APPLICATION NUMBER: 60/128,514  
 ; PRIOR FILING DATE: 1999-04-09  
 ; PRIOR APPLICATION NUMBER: 60/186,592  
 ; PRIOR FILING DATE: 2000-03-03  
 ; NUMBER OF SEQ ID NOS: 152  
 ; SOFTWARE: Curaseq1 version 0.1  
 ; SEQ ID NO 146  
 ; LENGTH: 142  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-189-940-146

Query Match 29.3%; Score 430.5; DB 16; Length 142;  
 Best Local Similarity 63.2%; Pred. No. 7.4e-32;  
 Matches 86; Conservative 12; Mismatches 37; Indels 1; Gaps 1;

QY 1 MFSGLTNCVLLLLQLLARSLEDGKVEGKNAYLPCSYTLPTSGTLVPMQWGKGFPCPW 60  
 DB 1 MFSHLPFCVLLLLLLLTSSSEYRRAVEGQNALPCFTTPAAGNLVPMQWGKACPV 60  
 QY 61 SDCNEHLRTBERNTYKSSRYQLKGLNKGVSLLIKNTVLDHGTGCCRIQPGIMN 120  
 DB 61 FECSNVVLRTERDVNW-TSRWMLNGDFRKGDVSLTIENTVLADSGIYCCRIQPGIMN 119  
 QY 121 DKLEIKLIDIKAKVT 136  
 DB 120 DEKFNKLVIKPGWT 135

Search completed: November 22, 2003, 05:56:14  
 Job time : 44.0052 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd..

OM protein - protein search, using sw model

Run on: November 22, 2003, 05:17:26 ; Search time 23.1753 Seconds  
(without alignments)  
1166.045 Million cell updates/sec

Title: US-10-004-633-10

Perfect score: 1468  
Sequence: 1 MFSGLTINCVLLILLQLLLAR.....VVEVENENXYCYVNSQOPS 281

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR.76:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	267.5	18.2	451	2	S71754
2	136.5	9.3	365	2	UC7780
3	131	8.9	503	2	UC5287
4	123	8.4	309	2	149503
5	120.5	8.2	403	2	152590
6	118.5	8.1	538	2	UC2457
7	116	7.9	459	2	A46254
8	111.5	7.6	247	2	A55717
9	106	7.2	186	2	161783
10	106	7.2	408	1	LKRT2
11	105	7.1	215	2	A57843
12	104	7.1	354	1	LKHU
13	104	7.1	509	2	JCS288
14	104	7.1	513	2	UC5288
15	103.5	7.1	354	1	S42938
16	102.5	7.0	457	2	A27449
17	102	6.9	457	1	RWMT4
18	102	6.9	457	2	T16525
19	101.5	6.9	458	1	RWHU74
20	101.5	6.9	1021	2	139207
21	101.5	6.9	1323	2	PN0568
22	101.5	6.9	1462	2	T42633
23	101	6.9	458	2	JC1509
24	101	6.9	521	2	S34338
25	100	6.8	364	2	A30521
26	98.5	6.7	346	2	S46993
27	98.5	6.7	354	1	S04243
28	98	6.7	562	2	G02426
29	97	6.6	3295	2	AE0074

30	96.5	6.6	1006	2	T31685	beta-galactosidase
31	96	6.5	345	2	I48780	Serai/Ep192 protei
32	95.5	6.5	349	2	A34815	carcinoembryonic a
33	95.5	6.5	1199	2	T23005	hypothetical prote
34	95	6.5	246	2	A47712	myelin/oligodendro
35	95	6.5	368	2	S17428	interleukin-1 rece
36	95	6.5	764	1	ORHUGS	secretory componen
37	95	6.5	967	2	S66852	hypothetical prote
38	95	6.5	1259	2	S36126	neural cell adhesi
39	94.5	6.4	1033	2	S19247	cell adhesion prot
40	94	6.4	272	2	C49209	outer surface prot
41	94	6.4	458	1	WMSR1	biliary glycoprote
42	94	6.4	521	2	JC1508	biliary glycoprote
43	94	6.4	523	2	I50478	neuroilin - goldfis
44	94	6.4	1257	1	A41060	neural cell adhesi
45	92.5	6.3	218	2	B47712	myelin/oligodendro

## ALIGNMENTS

## RESULT 1

S71754 cellular hepatitis A receptor HAVcr-1 precursor - green monkey

N/Alternate names: surface glycoprotein

C/Species: Cercopithecus aethiops (green monkey, grivet)

C/Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 05-Jun-1998

C/Accession: S71754

R/Kaplan, G.; Totenuka, A.; Thompson, P.; Akatsuka, T.; Moritsugu, Y.; Feinstein, S.M.

EMBO J. 15, 4282-4296, 1996

A/Title: Identification of a surface glycoprotein on African green monkey kidney cells a

A/Reference number: S71754; MUID:97015129; PMID:8661957

A/Accession: S71754

A/Molecule type: mRNA

A/Residues: 1-451 <Kap>

A/Cross-references: EMBL:X98252; NID:g1526573; PID:e247449; PID:g1526574

A/Experimental source: Kidney

C/Keywords: glycoprotein

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-451/Product: cellular hepatitis A receptor HAVcr-1 #status predicted <MAT>

Query Match	18.2%	Score 267.5;	DB 2;	Length 451;
Best Local Similarity	21.7%;	Pred. No. 1.86-14;		
Matches	97;	Conservative 51;	Mismatches 96;	Indels 203; Gaps 11;
QY	7	LNCVLLLOLLARSLEDGKVE--VGKNAVLPESYTLPTPSGTVPMCKGKGFPMQOCT	64	
DB	3	LQVILSLILHLADSVADSVVDGAGLSTLPCRY---NGAITSMCWNRGTCSVFSQP	58	
QY	65	NEILRTDERNVTYQSSRYOLKGDVSLIKNTVLDHGYCCRIOPGLMNDKUL	124	
DB	59	DGIWVNGTHVTRKTRKRLKGLNLSRRVSLTIANAVSDGILYCRVAKHSGFNDWKI	118	
QY	125	ELKLDIRAKVT	136	
DB	119	TISLAKIGPRNTPIVATVTRSTVPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPT	178	
QY	137	-----	136	
DB	179	LPTTTVPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPT	238	
QY	137	-----	136	
DB	239	TTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPT	298	
QY	137	-----	136	
DB	299	PVATSPSPQAEHNPVTLGATRTQPTSPSLVYTTDGSPTVRESSGLWNNNOTQSP	358	
QY	179	WADIKSGERTIR--AIHIGVGSAGTLALIGVLLKRYSKKKKLSLTLTANL	236	
DB	359	-----EHSPOWNTTBGIYAGVCISV-LVLLAVGVIAKKYFP-KKEIQQLS-VSPSNH	410	

QY 237 PPGGLANAGAVRIRSEENITYTEENY 263  
 Db 411 QPKTLQNAVKEVHADNITY-IENNL 436

## RESULT 2

coxackie- and adenovirus receptor - bovine  
 Jc7780

C:Species: Bos primigenius taurus (cattle)  
 C:Date: 02-Apr-2002 #sequence\_revision 02-Apr-2002 #text\_change 02-Apr-2002

C:Accession: Jc7780

R:Thoclem, I.; Keyaerts, E.; Lindberg, M.; Van Ranst, M.  
 Biochem. Biophys. Res. Commun. 288, 805-808, 2001

A:Title: Characterization of a cDNA encoding the bovine coxackie and adenovirus recepto  
 A:Reference number: Jc7780

A:Contents: Liver

A:Accession: Jc7780

A:Molecule type: mRNA

A:Residues: 1-365 <THO>

A:Cross-references: GB:AY033651

C:Comment: This protein serves as the primary adenoviral attachment site on bovine cells

Query Match 9.3%; Score 136.5; DB 2; Length 365;  
 Best Local Similarity 22.5%; Pred. No. 0.0011;  
 Matches 61; Conservative 44; Mismatches 99; Indels 67; Gaps 11;

QY 12 LLLQLLARSLEB--GYRVEV-----GKNATLPSCSYTLPTSGTLVPMCKGKFCP 59

Db 3 LLLRFLLCGVADFTGSLITTPPEQMIERAKGETAVLPCKFTLG-----PEDQGLDIE 56

QY 60 WSOCTEELRTDERNVTYQKSSRY-----QLKG-----DLKGDVSLIKNVTLDDHG 107

Db 57 WLSPADNOKVDVILYSGDKTYDYODLKRVAFTSNDLKSGDASINVTNLQLSDIG 116

QY 108 TYCCRI-QPPGLMNDKLEIKLIDKAA-----KYTPAQTA---HGD 144

Db 117 TYQCKYKKAQVGN-KKIQTLVLRKSGRCVYDGESEIENDFKLCPEKESGLPIRYEM 175

QY 145 STTASRITLTENGETQTQTLVTHNNNGTKISTWDEIKDGETRTAIHIGV----- 198

Db 176 QKLSQDKQKPTSWLPMTSPVISVKNASAEVSGTYCTVARNRVSQDCLRLDVPSPNR 235

QY 199 ----GVSAGLTLALIIIVLILKWSCKKK 224

Db 236 AGTIAGAVIGTLALVIALIV--FCHKKR 264

## RESULT 3

SHP substrate-1 protein - human

Jc5287

C:Species: Homo sapiens (man)  
 C:Date: 16-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 20-Jun-2000

C:Accession: Jc5287

R:Yamato, T.; Mazonaki, T.; Amano, K.; Matsuda, Y.; Takahashi, N.; Ochi, F.; Fujioka, Y.;  
 Biochem. Biophys. Res. Commun. 231, 61-67, 1997

A:Title: Mouse and human SHPS-1: Molecular cloning of cDNAs and chromosomal localization  
 A:Reference number: Jc5287; MUID:97223399; PMID:9070220

A:Contents: Brain

A:Accession: Jc5287

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-503 <YAM>

A:Cross-references: DDBJ:D86043; NID:g1864010; PIN:BA12974.1; PID:g1864011  
 C:Comment: This protein is a glycosylated receptor-like protein and plays a role in cell  
 acts as a docking protein and induce translocation of SHP-2 from the cytosol to the pl

C:Genetics:

A:Gene: shps-1

A:Map position: 20p13

Query Match 8.9%; Score 131; DB 2; Length 503;  
 Best Local Similarity 22.5%; Pred. No. 0.0047;  
 Matches 67; Conservative 45; Mismatches 124; Indels 62; Gaps 13;

QY 13 LLLQLLARSLEB-----EDGYVEVGNAYLPCSYT--LPTSGTLVPMCKG 54  
 Db 14 LLLQLLARSLEB-----EDGYVEVGNAYLPCSYT--LPTSGTLVPMCKG 54  
 QY 55 KGFPCWQCTNELRTDE-----RNVTYQKSSRYQKDLNKGDVSLIKNVTLDDHGY 109  
 Db 70 RQAGP-----ARLITNQGKGFPRVTVSESTKRE-----NMDSISISNITPDAQTY 119  
 QY 110 CCRIFPGLMNDKLE---LKLIDKAAKVTPTAQAHDSTTASPR---TLTENGSET 162  
 Db 120 YC-VKFRKSGSPTEFKSGAGTELAVRAKPSAP--VSGPAAATPQHTVSEFCSEHGFSP 176  
 QY 163 QTLVTLHNNNGTKISTWDEIKDGETRTAIHIGVSGAGTLTLLIIGVILKWSCKK 222  
 Db 177 RQTLKWRKNGELSDPQTNVDPGEVSYSISHTAKV-----VLTREDVHS 223  
 QY 223 KLLSSLSLITLANLPGLANAGAVRIRSEENITYTEENYEVENSNEYCVNSQOP 280  
 Db 224 QVICAVAVTLQGDPLRGTAFLSER-IKVPITLFTYQGPV-RAENQVAVTCQVRKFP 279

## RESULT 4

I49503

B-lymphocyte activation antigen 7 precursor - mouse

N:Alternate names: MB7-2

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 29-Sep-1999

C:Accession: I49503; S17291; I49521

R:Selyakumar, A.; White, P.C.; Dupont, B.  
 Immunogenetics 38, 292-295, 1993

A:Title: Genomic organization of the murine B-lymphocyte activation antigen B7.  
 A:Reference number: I49503; MUID:93307789; PMID:7686531

A:Accession: I49503

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-309 <RES>

A:Cross-references: GB:I12589; NID:9293299; PID:AAA37240.1; PID:9293301

R:Freeman, G.J.; Gray, G.S.; Gimmi, C.D.; Lombard, D.B.; Zhou, L.O.; White, M.; Fingerot  
 J. Exp. Med. 174, 625-631, 1991

A:Title: Structure, expression, and T cell costimulatory activity of the murine homologu  
 A:Reference number: S17291; MUID:91341422; PMID:1714935

A:Accession: S17291

A:Molecule type: mRNA

A:Residues: 1-274, 'R', 279-309 <PRE>

A:Cross-references: EMBL:X60958; NID:950111; PID:CAA42291.1; PID:950112

R:Inobe, M.; Linsley, P.S.; Ledbetter, J.A.; Nagai, Y.; Tanakoshi, M.; Ueda, T.  
 Biochem. Biophys. Res. Commun. 200, 443-449, 1994

A:Title: Identification of an alternatively spliced form of the murine homologue of B7.  
 A:Reference number: I49521; MUID:94220123; PMID:7513163

A:Accession: I49521

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-143, 238-274, 'R', 279-309 <RES>

A:Cross-references: GB:D16220; NID:9505118; PID:BA03748.1; PID:994769

C:Genetics:

A:Gene: B7

A:Introns: 37/1, 143/1, 237/1, 275/1

C:Superfamily: B-lymphocyte restricted antigen B7

C:Keywords: alternative splicing

Query Match 8.4%; Score 123; DB 2; Length 309;  
 Best Local Similarity 24.2%; Pred. No. 0.012; Indels 90; Gaps 12;

Matches 67; Conservative 31; Mismatches 89; Indels 90; Gaps 12;

QY 11 LLLQLLARSLEB-----SLDGYKVEVGNAYLPCSYTPTSG-----TLVPM 51

Db 21 LLLRFLIRLSQVSDVDQSKSVKXKVLPCRYNPHDESDRITVWQKRDVLSV 80

QY 52 CWKGFPCWQCTNELRTDERNVTYQKSSRYQLKGLNKGDVSLIKNVTLDDHGYCC 111

Db 81 IAGK-LKWPEYKN--RTLYDNTTY-----SLIIGLVLSDRGTVSC 119

QY 112 RIQFP--GLMNDKLEL-KLDIKAAKVTPTAQAHDSTTASPR----- 151

Db 120 VQKKERGTVEVKHLATLVKLSIKADFTSPNITESGNPSADTKRTCFASGGFPKPRFSWL 179  
Qy 152 -----TLTERNGSEFQTLVTLHNNNGTK-----IS---TWABEIK 184  
Db 180 ENGELPOINTTISQDPESELTYTSSQDPEFTTRNHTIKLKYGDAAHVSDFTWEKPE 239  
Qy 185 DSGETIRTAIHIGVVSAGLTALILIGVILKMYSC 221  
Db 240 DPPSKNTLVLFAGAGFAGVITVVI--VVIKCF-CK 273

## RESULT 5

152590  
m33-B isoform - mouse  
C:Species: Mus sp. (mouse)  
C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: 152590  
R:Chilian, E.Z.; Beverley, P.C.; Young, B.D.; Watt, S.M.  
Blood 83, 3188-3198, 1994  
A:Title: Molecular cloning of two isoforms of the murine homolog of the myeloid CD33 and  
A:Reference number: 152590; MUID:94250900; PMID:8193354  
A:Accession: 152590  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-403 <RES>  
A:Cross-references: GB:871345; NID:9551352; PIDN:AA830842.1; PID:9551353

Query Match 8.2%; Score 120.5; DB 2; Length 403;  
Best Local Similarity 23.7%; Pred. No. 0.026;  
Matches 65; Conservative 32; Mismatches 116; Indels 61; Gaps 8;

Qy 10 VLLILQLLARSLE-----DGYKVEGKNAYLPCSYTLF-----TSGTLVVMCMGKGCPC 59  
Db 7 LFLCAGSLADLEFQVAPSVTVEBSLCVHPSPVFPYPIKILGLGVGSGMIRKGV-- 64  
Qy 60 WSQCTNELRTDERNTYQKSS--RYQKGLDINKGVSLIKNTVLDHGTGCRIOFP 117  
Db 65 -SLHEDSPVATSDRQLVQKATQGRFQLGDPQKHDCSLFRDQKNDTGMVFRVAREP 123  
Qy 118 LM-NDKLELKLIDKAAKTPAQTAGDSTTASPRITL----- 154  
Db 124 FVRYSYKKSQSLVHTLSRTPDIIIPGLTAGYPSNLTGCVPAECGTPPTFSWMSTA 183  
Qy 155 -----TENNGSETQTLVTLHNNNGTKISTMADEIKDSETRTAIHIGVVSAG----- 203  
Db 184 LTSLSRTTSSVLTFTPOPDHGTCL-TCLVTSGAGVTVERTIQLNVTBKSGQMRRLV 242  
Qy 204 -----LTLALIGVILKMYSCKKKLS 226  
Db 243 LVAVGEATVKLLIIGLCVFLIWMCRKRTKLS 276

## RESULT 6

JC2457  
vascular cell adhesion protein - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 15-Feb-1995 #sequence\_revision 05-Apr-1995 #text\_change 08-Oct-1999  
C:Accession: JC2457  
R:Taang, Y.T.M.; Haskard, D.O.; Robinson, M.K.  
Biochem. Biophys. Res. Commun. 201, 805-812, 1994  
A:Title: Cloning and expression kinetics of porcine vascular cell adhesion molecule.  
A:Reference number: JC2457; MUID:94271236; PMID:7516159  
A:Accession: JC2457  
A:Molecule type: mRNA  
A:Residues: 1-538 <TSA>  
A:Cross-references: EMBL:U08351; NID:9474382; PIDN:AAA21542.1; PID:9474383  
C:Keywords: glycoprotein; transmembrane protein  
F:497-517/Domain: transmembrane #stratus predicted <TM>  
F:75,157,271,330,360/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 8.1%; Score 118.5; DB 2; Length 538;  
Best Local Similarity 25.1%; Pred. No. 0.056;

Matches 63; Conservative 42; Mismatches 81; Indels 65; Gaps 15;  
Qy 63 CTNELRTDERNTYQKSSRYQKGLDINKGVSLIKNTVLDHGTGCRIOF 115  
Db 244 CTSEGLPAPD--ISWSK-----LDNGDQQLSGNATLTLINAMESGIYECGVN 293  
Qy 116 PGLMNDKLELKLID-----KAAKTPAQTAGDSTT-----ASPRITLITE--RNGS 160  
Db 294 PVGTNRKEVELTYQVARDTTISVNPSTLEEGSSVMTCSDGFPAPKILMSKLRDGN 353  
Qy 161 ETQTLVTLHNNNGTKISTMADEIKDSETRTAIHIGVVS-AGLT--LALITGV----- 212  
Db 354 -----LEPLSEN--TTLTLTSTKMDSG-----IYVCEGINQAGINKKEVELIIQAAPKD 401  
Qy 213 LILKWSCKKKKLSLSLI--TLANLPPGGL-----ANAGAVRIRSENYTYT-----E 259  
Db 402 LQLPAPSESVKEDTVITCTCGNVPTLILKKKAFEDTVLKSTDGAITTHRAFLAD 461  
Qy 260 ENVEVENSNE 270  
Db 462 AGVVECESKNE 472

## RESULT 7

A46254  
CD4 precursor - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jan-2000  
C:Accession: A46254  
R:Hague, B.F.; Sawasaki, S.; Brown, T.J.; Lee, K.; Recker, D.P.; Kindt, T.J.  
Proc. Natl. Acad. Sci. U.S.A. 89, 7963-7967, 1992  
A:Title: CD4 and its role in infection of rabbit cell lines by human immunodeficiency vi  
A:Reference number: A46254; MUID:92390370; PMID:1518821  
A:Accession: A46254  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-459 <HAG>  
A:Cross-references: GB:M92840; NID:g164871; PIDN:AAA31198.1; PID:g164872  
A>Note: sequence extracted from NCBI backbone (NCBI:112732; NCBI:P:112733)  
C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
F:322-372/Domain: immunoglobulin homology <IM>

Query Match 7.9%; Score 116; DB 2; Length 459;  
Best Local Similarity 26.8%; Pred. No. 0.074;  
Matches 45; Conservative 21; Mismatches 54; Indels 48; Gaps 7;

Qy 9 CVLILQLLARSLEDGYKVEGKNAYLPCSYTLPTSGTLVPMCMGKGCPCWSQCTNEL 68  
Db 9 CLILVLPALIPATWGTQVTRGK-----AGALVELP-----CQSSQKRNVSF 51  
Qy 69 RTDERNTY-----QKSSRYQKGD-----LNKGVSLIKNTVLDHGT 109  
Db 52 NWRKANQVKILGNQSSSSFWLKGNSPLSNRVESKKNMWDQGSFPLVITDLRMDSGTY 111  
Qy 110 CCRIOFPGLMNDKLELKLIDKAAKTP-AGTADGDTTASPRITLITE 156  
Db 112 ICEV-----GDKMEVELVFRITANPNTRLHGQSL-----TLTLE 148

## RESULT 8

A55717  
myelin/oligodendrocyte glycoprotein precursor - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 08-Oct-1999  
C:Accession: A55717; C47712  
R:Daubae, P.; Pham-Dinh, D.; Dautigny, A.  
Genomics 23, 36-41, 1994  
A:Title: Structure and polymorphism of the mouse myelin/oligodendrocyte glycoprotein gen  
A:Reference number: A55717; MUID:95130110; PMID:7829100  
A:Accession: A55717  
A:Molecule type: DNA  
A:Residues: 1-247 <DAU>  
A:Cross-references: GB:L29498

R.Pharm.Dinh, D.; Martel, M.G.; Nusbaum, J.L.; Roussel, G.; Pontarotti, P.; Roeckel, N.;  
Proc. Natl. Acad. Sci. U.S.A. 90, 7990-7994, 1993  
A/Title: Myelin/Oligodendrocyte glycoprotein is a member of a subset of the immunoglobul  
A/Reference number: A47712; MUID:93376728; PMID:8367453  
A/Accession: C47712  
A/Molecule type: mRNA  
A/Residues: 30-95, 'E', 97-247 <PHA>  
A/Cross-references: GB:L20942; NID:g399588; P1DN:AAA03180.1; PID:g399589  
C/Genetics:  
A/Genes: MOG  
A/Map position: 17  
A/Note: encoded within the MHC  
C/Function:  
A/Description: may be involved in lipid interaction; may be involved in cell-cell commu  
C/Keywords: glycoprotein; myelin; transmembrane protein  
F/1-23/Domain: signal sequence #status predicted <SIG>  
F/30-247/Product: myelin/oligodendrocyte glycoprotein #status predicted <MAT>  
F/151-179/Domain: transmembrane #status predicted <TM1>  
F/204-229/Domain: transmembrane #status predicted <TM2>  
F/60/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.6%; Score 111.5; DB 2; Length 247;  
Best Local Similarity 26.6%; Pred. No. 0.08;  
Matches 37; Conservative 28; Mismatches 53; Indels 21; Gaps 5;

QY 8 NCYLLLLQLLA-----RSLDGYKE--VGNKAYLPCTSLPTSGTIVPMCKGK 55  
DB 11 SCPLSLLLLLQLQSCSYAGQFVIGPGYPIRALVDEDELPRISPGKNAATGMEVGYR 70  
QY 56 GFCFMSQCTNELRLTDERANTYOKSSRYQ----LKGDINKGVSLLIKNVTLDHGTGTC 110  
DB 71 S--PFRVYH--LYRNGKQDAEQAPERYKGRTELLKETTISEGKVTIRIQNVRPSDGGYT 126  
QY 111 CRIQPGKLMNDKKLEKLD 129  
DB 127 CFEFRDHSYQBEAMELKVE 145

RESULT 9  
161783  
sodium channel beta 2 subunit - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 05-Nov-1999  
C/Accession: I61783  
R/Isom, L.L.; Ragsdale, D.S.; DeJongh, K.S.; Westebroek, R.E.; Reber, B.F.X.; Schuer,  
Cell 83, 433-442, 1995  
A/Title: Structure and Function of the Beta 2 Subunit of Brain Sodium Channels,  
A/Reference number: A57843; MUID:96067641; PMID:8521473  
A/Accession: I61783  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-186 <RES>  
A/Cross-references: EMBL:U37147; NID:g1086498; P1DN:AAB60506.1; PID:g1086499  
C/Genetics:  
A/Genes: SCN2  
A/Introns: 50/3; 120/2

Query Match 7.2%; Score 106; DB 2; Length 186;  
Best Local Similarity 23.2%; Pred. No. 0.16;  
Matches 52; Conservative 29; Mismatches 77; Indels 66; Gaps 8;

QY 31 GKNAYLPCSYT--LPTSGTIVPMCKGKCPMSQCTNELRLTDERANTYOKSSRY----Q 84  
DB 14 GSDTRLPTCPNSCYTVNHHKPSLNTVYQEC--SNCESENVLOFRMKIINLKERFEDRAVE 71  
QY 85 LKGDINKGVSLLIKNVTLDHGTGTCRIQFPGIAMDKKLELTDIKAAKVTPTAQTADHD 144  
DB 72 PEGNPEKIVSVATLKRVQLEDEGIVNCYITNPDRRHGKHTLVQV--LLEPPER----D 126  
QY 145 STTASPRTLTTERNSEGTQTLVTLHNNNGTKISTWADKIDSGETIRTAIHIGVVSAGI 204  
DB 127 STVA-----VIVGASVGG 139

QY 205 TIALITGVILKWSYSC---KKKKSLSLITLNLAPPGGLANA 244  
DB 140 FLAAVVI--LVLMVVCVRRKKEQKLTSTDLKTEERKGTGDSGA 181

RESULT 10  
LKRT2  
proteoglycan link protein 2 precursor - rat  
N/Alternate names: cartilage link protein  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 13-Aug-1986 #sequence\_revision 24-Oct-1997 #text\_change 13-Nov-1998  
C/Accession: A28654; A24880; A02869  
R/Rhodes, C.; Doege, K.; Sasaki, M.; Yamada, Y.  
J. Biol. Chem. 263, 6063-6067, 1988  
A/Title: Alternative splicing generates two different mRNA species for rat link protein.  
A/Reference number: A28654; MUID:88198139; PMID:2452158  
A/Accession: A28654  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-408 <RHO>  
R/Doege, K.; Hassell, J.R.; Caterson, B.; Yamada, Y.  
Proc. Natl. Acad. Sci. U.S.A. 83, 3761-3765, 1986  
A/Title: Link protein cDNA sequence reveals a tandemly repeated protein structure  
A/Reference number: A24880; MUID:86233314; PMID:3459153  
A/Accession: A24880  
A/Molecule type: mRNA  
A/Residues: 180-408 <DOE>  
R/Name, P.J.; Christner, J.R.; Baker, J.R.  
J. Biol. Chem. 261, 3519-3535, 1986  
A/Title: The primary structure of link protein from rat chondrosarcoma proteoglycan aggr  
A/Reference number: A02869; MUID:86140139; PMID:2419334  
A/Accession: A02869  
A/Molecule type: protein  
A/Residues: 16-33, 'A', 89-375, 'W', 377-408 <NEA>  
C/Comment: This protein was extracted from rat chondrosarcoma.  
C/Comment: Residues 259-277 and 358-376 (approximately) form epitopes for the species-n  
C/Comment: Link proteins interact with and stabilize aggregates of hyaluronic acid and c  
C/Suprafamily: proteoglycan link protein; immunoglobulin homology; link protein repeat  
C/Keywords: cartilage; chondroitin sulfate proteoglycan; duplication; glycoprotein; hyal  
F/1-15/Domain: signal sequence #status predicted <SIG>  
F/16-408/Product: proteoglycan link protein 2 #status experimental <MAT>  
F/108-195/Domain: immunoglobulin homology <IMM>  
F/230-307/Domain: link protein repeat homology <LNK1>  
F/328-404/Domain: link protein repeat homology <LNK2>  
F/110/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F/115-193, 235-306, 259-280, 333-403, 358-379/Disulfide bonds: #status experimental

Query Match 7.2%; Score 106; DB 1; Length 408;  
Best Local Similarity 29.9%; Pred. No. 0.43;  
Matches 40; Conservative 21; Mismatches 39; Indels 34; Gaps 8;

QY 20 RSLDGYKEV-----GKNAYLPCSYTLPSTGTVPMCKGKF---CWSQCT 64  
DB 86 RTTENGPRLLVBAEQAKVFSHRGQVNTLPCKRYRD-----PTAGSGIHKIRIKWTXLT 139  
QY 65 NELRLTDERANTY--YOKSS-----RYQLKGLINKGVSLLIKNVTLDHGTGTCRIQFP 116  
DB 140 SYLREVDVDFVSMGHHKHTYGGYGVFLKGG--SNDASLITLTDLLEDRGRKCEV--IE 197  
QY 117 GLMNDK--KLELK 127  
DB 198 GLEDTAVVALELQ 211

RESULT 11  
A57843  
sodium channel beta 2 subunit - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C/Accession: A57843  
R/Isom, L.L.; Ragsdale, D.S.; DeJongh, K.S.; Westebroek, R.E.; Reber, B.F.X.; Schuer,  
Cell 83, 433-442, 1995  
A/Title: Structure and Function of the Beta 2 Subunit of Brain Sodium Channels, a Trans

A:Reference number: A57843; MUID:96067641; PMID:8521473  
 A:Accession: A57843  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1215 <RES>  
 A:Cross-references: EMBL:U37026; NID:91086496; PIDN:AA53967.1; PID:91086497  
 C:Genetics:  
 A:Gene: SCNB2

Query Match 7.2%; Score 105; DB 2; Length 215;  
 Best Local Similarity 23.2%; Pred. No. 0.23;  
 Matches 52; Conservative 28; Mismatches 78; Indels 66; Gaps 8;

QY 31 GKNAVPCSYT--LPTSGTLVPMCMGKGFPCMSQCTNELRTDERNVYOKSSRY---Q 84  
 DB 43 GSDTRLPCTNSCYVHHKQPSLMTWTOEC--SNCSSEMLQFMKINKLTERFGDRI 100  
 QY 85 LKCDLNGKDVSLIKNVTLDDHGTCCRIQFGLMNDKLELKDIAKAVTPAQTAGD 144  
 DB 101 FSGNPSKYDVSVTLKNVQLEDEGIYNCYITNPDRHGHGKIYLOV--LLEVPPEP---D 155  
 QY 145 STIASPRTLTEENGSOTLVTLHNNNGKISWMADEIIDSGETITTAHIGVSAQL 204  
 DB 156 STVA-----VIVGASVVG 168  
 QY 205 TLLAIIGVLIKWYSC---KKKLSLSLITLTLANLPPGLANA 244  
 DB 169 FLAVVI--LVLMVVKCVRRKKEQKLTSDDLKTEBEGKTDEGNA 210

## RESULT 12

LKHU  
 proteoglycan link protein precursor [validated] - human

C:Species: Homo sapiens (man)  
 C>Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 08-Dec-2000  
 C:Accession: S14914; S04244; S03868; A36308; S14926; S09309; S08041  
 R:Dudhia, J.; Hardingham, T.E.  
 Nucleic Acids Res. 18, 1292, 1990  
 A>Title: The primary structure of human cartilage link protein.  
 A:Reference number: S14914; MUID:90206798; PMID:2320422  
 A:Accession: S14914  
 A:Molecule type: mRNA  
 A:Residues: 1-354 <DUD>  
 A:Cross-references: EMBL:X17405; NID:9463246; PIDN:CAA35462.1; PID:934378  
 A>Note: the authors translated the codon GAT for residue 93 as Ala  
 R:Dudhia, J.; Hardingham, T.E.  
 J. Mol. Biol. 206, 749-753, 1989  
 A>Title: Appendix. Isolation and sequence of cDNA clones for pig and human cartilage link  
 A:Reference number: S04243  
 A:Accession: S04243  
 A:Molecule type: mRNA  
 A:Residues: 223-354 <DUD>  
 A:Cross-references: EMBL:Y00166  
 A>Note: the authors translated the codon CTG for residue 264 as Arg and TTG for residue  
 R:Nguyen, Q.; Murphy, G.; Roughley, P.J.; Mort, J.S.  
 Biochem. J. 259, 61-67, 1989  
 A>Title: Degradation of proteoglycan aggregate by a cartilage metalloproteinase. Evidence  
 A:Reference number: S03868; MUID:89246328; PMID:2719651  
 A:Accession: S03868  
 A:Molecule type: protein  
 A:Residues: 16-35 <NGU>  
 R:Osborne-Lawrence, S.L.; Sinclair, A.K.; Hicks, R.C.; Lacey, S.W.; Eddy Jr., R.L.; Byer  
 Genomics 8, 562-567, 1990  
 A>Title: Complete amino acid sequence of human cartilage link protein (CRTL1) deduced fr  
 A:Reference number: A36308; MUID:91139126; PMID:2286376  
 A:Accession: A36308  
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
 A:Molecule type: mRNA  
 A:Residues: 1-354 <OSB>  
 R:Dudhia, J.; Hardingham, T.E.  
 Nucleic Acids Res. 18, 2214, 1990  
 A>Title: The primary structure of human cartilage link protein.

A:Reference number: S14926; MUID:90245703; PMID:2336413  
 A:Accession: S14926  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-92, 'A', '94-354 <NUC>  
 R:Nguyen, Q.; Mort, J.S.; Roughley, P.J.  
 Biochem. J. 266, 569-573, 1990  
 A>Title: Cartilage proteoglycan aggregate is degraded more extensively by cathepsin L th  
 A:Reference number: S09309; MUID:90197639; PMID:2317204  
 A:Accession: S09309  
 A:Molecule type: protein  
 A:Residues: 16-38, 'X', '40-55, 57-60, 62-65, 'X', '67, 'X', '69-80 <NG2>  
 C:Genetics:

A:Gene: GDB:CRTL1  
 A:Cross-references: GDB:125232; OMIM:115435  
 A:Map position: 5q13-5q14  
 C:Superfamily: proteoglycan link protein; immunoglobulin homology; link protein repeat  
 C:Keywords: cartilage; duplication; extracellular matrix; glycoprotein  
 F:1-15/Domain: signal sequence #status predicted <SIG>  
 F:16-354/Product: proteoglycan link protein #status experimental <MAT>  
 F:54-141/Domain: immunoglobulin homology <IMM>  
 F:176-253/Domain: link protein repeat homology <LNK1>  
 F:274-350/Domain: link protein repeat homology <LNK2>  
 F:21,56/Binding site: carbohydrate (asn) (covalent) #status predicted  
 F:61-139/Dissulfide bonds: #status predicted

Query Match 7.1%; Score 104; DB 1; Length 354;  
 Best Local Similarity 24.2%; Pred. No. 0.53;  
 Matches 39; Conservative 27; Mismatches 47; Indels 48; Gaps 8;

QY 10 VLLQLQLLARSLEDEYKVE-----VGRNAVPCSYTL 42  
 DB 6 LVLVISICVADHDLSNDVTLDDHRAIHQAENGPHLVEAQAQVFSRGNAVTLPCPFYR 65  
 QY 43 PTSGTLVPMCMGKGF---CPMSQCTNELRTDERNVY--YOKSS-----RYLKGDLN 90  
 DB 66 D-----PTAFSGGSHIRIKRKLTSDYLKEVDVFWSMGYHKKTYGGYGRVFLKGG-S 118  
 QY 91 KGDVSLIKNVTLDDHGTCCRIQFGLMNDKLELKDIAKAVTPAQTAGD 131  
 DB 119 DSDASLVITDLTLEDYGRKCEV--IRGL-EDDTVVALDLQ 157

## RESULT 13

JCS288  
 SHP substrate-1 protein, 509 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 16-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 20-Jun-2000

C:Accession: JCS288

R:Yamao, T.; Matozaki, T.; Amano, K.; Matsuda, Y.; Takahashi, N.; Ochi, F.; Fujioaka, Y.;

Biochem. Biophys. Res. Commun. 231, 61-67, 1997

A>Title: Mouse and human SHP-1: Molecular cloning of cDNAs and chromosomal localization

A:Contents: Brain

A:Accession: JCS288

A:Molecule type: mRNA

A:Residues: 1-509 <YAM>

A:Cross-references: DDBJ:D87967; NID:91864012; PIDN:BA13520.1; PID:91864013

C:Comment: This protein is a glycosylated receptor-like protein and plays a role in cell

acts as a docking protein and induce translocation of SHP-2 from the cytosol to the pla

C:Genetics:

A:Gene: shps-1

A:Map position: 2

Query Match 7.1%; Score 104; DB 2; Length 509;  
 Best Local Similarity 20.4%; Pred. No. 0.84;  
 Matches 68; Conservative 39; Mismatches 97; Indels 130; Gaps 14;

QY 12 LLLQLQLLARSLEDEYKVEKNAVPCSYT--LPTSGTLVPMCMG 54  
 DB 14 LLLQLQLLARSLEDEYKVEKNAVPCSYT--LPTSGTLVPMCMG 54  
 QY 55 KGFPCMSQCTNELRTDERNVYOKSSRYOLKGD-----LNKGDVSLIKNV 101

```

Db      70 RG-----VQSRLIYSFTGEHPPRTVNSDATKRNMDPSIRISNV 111
Qy      102 TLDDHGYCCRIQFPGIANDKLEKLD-----IKAKVTPAQT----- 140
Db      112 TPEDAGTYCC-VKFGKGPSEPDTEIQSGGSTEYVILAKSPPEVSGPADRGIPDQKVNPT 170
Qy      141 --AHGDSSTTASPRITLITE--RNGSETQTLVTLHNNNGTKISTWABEIKDSGETIRTALHI 196
Db      171 CKSHG---FSPRNITLKWFKDGOELHLETTVNPBGKNVS-----VNISSTVR----- 215
Qy      197 GVGVSAGLTLALIIIGVILIKWYSCKKKLSLITLANLPFGGLAN-----AGAVRI 249
Db      216 -----VLIANSMDVHSKVICVAHITLDRSPLRGIANLSNFIKVSPTVKV 259
Qy      250 RSE-----ENIYTIENNVYEVNSN 269
Db      260 TQGSPTSMNQVNLTCRAERFYPEDLQIWLKNGN 293

```

## RESULT 14

```

JCS289
SHP substrate-1 protein, 513 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 16-Apr-1997 #sequence_revision 09-May-1997 #text_change 20-Jun-2000
C/Accession: JCS289
R/Yamao, T.; Maizumi, T.; Amano, K.; Matsumura, Y.; Takahashi, N.; Ochi, F.; Fujioaka, Y.;
Biochem. Biophys. Res. Commun. 221, 61-67, 1997
A/Title: Mouse and human SHP-1: Molecular cloning of cDNAs and chromosomal localization
A/Reference number: JCS289; MUID:97223399; PMID:97070220
A/Contents: Bzrain
A/Accession: JCS289
A/Molecule type: mRNA
A/Residues: 1-513 <YAM>
A/Cross-references: DBJ:087968; NID:91864014; PIDN:BA13521.1; PID:91864015
C/Comment: This protein is a glycosylated receptor-like protein and plays a role in cell
acts as a docking protein and induce translocation of SHP-2 from the cytosol to the pla
C/Genetics:
A/Gene: shp-1
A/Map position: 2

```

```

Query Match      7.1%; Score 104; DB 2; Length 513;
Best Local Similarity 20.4%; Pred. No. 0.85;
Matches 68; Conservative 39; Mismatches 97; Indels 130; Gaps 14;

Qy      12 LILQLLARSL-----EDGYKVEGKNAYLPQSYT--LFTSGTLVPMCWG 54
Db      14 LILCLLSASCFCTGVTKELKQPEKSVSAAGDSVLANCTLSLPVG---PIKMY 69
Qy      55 KGFQWQSCTNELLTDERNVYQKSSRYQLKGD-----LNKGVSLIKNV 101
Db      70 RG-----VQSRLIYSFTGEHPPRTVNSDATKRNMDPSIRISNV 111
Qy      102 TLDDHGYCCRIQFPGIANDKLEKLD-----IKAKVTPAQT----- 140
Db      112 TPEDAGTYCC-VKFGKGPSEPDTEIQSGGSTEYVILAKSPPEVSGPADRGIPDQKVNPT 170
Qy      141 --AHGDSSTTASPRITLITE--RNGSETQTLVTLHNNNGTKISTWABEIKDSGETIRTALHI 196
Db      171 CKSHG---FSPRNITLKWFKDGOELHLETTVNPBGKNVS-----VNISSTVR----- 215
Qy      197 GVGVSAGLTLALIIIGVILIKWYSCKKKLSLITLANLPFGGLAN-----AGAVRI 249
Db      216 -----VLIANSMDVHSKVICVAHITLDRSPLRGIANLSNFIKVSPTVKV 259
Qy      250 RSE-----ENIYTIENNVYEVNSN 269
Db      260 TQGSPTSMNQVNLTCRAERFYPEDLQIWLKNGN 293

```

## RESULT 15

```

S42938
proteoglycan link protein precursor - horse

```

```

N/Alternate names: Cartilage link protein
C/Species: Equus caballus (domestic horse)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 22-Jun-1999
C/Accession: S42938
R/Dublia, J.; Platt, D.
submitted to the EMBL Data Library, March 1994
A/Reference number: S42938
A/Accession: S42938
A/Molecule type: mRNA
A/Residues: 1-354 <DDU>
A/Cross-references: EMBL:X78077; NID:9459438; PIDN:CA454987.1; PID:9459439
C/Superfamily: proteoglycan link protein; immunoglobulin homology; link protein repeat h
C/Keywords: cartilage; duplication; extracellular matrix; glycoprotein
F/1-15/Domain: signal sequence #status predicted <SIG>
F/16-354/Product: proteoglycan link protein #status predicted <MAT>
F/54-141/Domain: immunoglobulin homology <IMM>
F/176-253/Domain: link protein repeat homology <LINK1>
F/274-350/Domain: link protein repeat homology <LINK2>
F/21-56/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/61-139/Disulfide bonds: #status predicted

```

```

Query Match      7.1%; Score 103.5; DB 1; Length 354;
Best Local Similarity 30.1%; Pred. No. 0.58;
Matches 34; Conservative 21; Mismatches 37; Indels 21; Gaps 7;

Qy      31 GKNAYLPQSYTLFTSGTLVPMCWGKGF-----CPMSQCTNELRTDERNVY--YQKSS--- 81
Db      54 GGNVTLPPKFLRD-----PTAFSGTTHKIRIKWTKLSDYLKEVDVFSWGMGHHKTYGG 107
Qy      82 ---RYQLGDNKGPVSLIKNVTLDDHGYCCRIQFPGIANDKLEKLDIK 131
Db      108 YQGRVFLKGG--SDNDASLVITDITDDYGRYKCEV--IGLEDDTV--VALDIQ 157

```

Search completed: November 22, 2003, 05:44:23  
Job time : 27.1753 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 01:50:35 ; Search time 13.5189 Seconds  
(without alignments)  
977.484 Million cell updates/sec

Title: US-10-004-633-10

Perfect score: 1468  
Sequence: 1 MFSGLTINCVLLLLQLLLAR.....VVEVENSNEYCYVNSQOPS 281

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157	10.7	319	A33_HUMAN	Q99795 homo sapien
2	147.5	10.0	365	CD4_MOUSE	P78310 homo sapien
3	134	9.1	398	SRB1_HUMAN	O00241 homo sapien
4	131	8.9	503	SRB1_HUMAN	P78324 h protein-t
5	129	8.8	387	SRB2_HUMAN	O99148 homo sapien
6	128.5	8.8	499	SRB2_HUMAN	O99124 homo sapien
7	126.5	8.6	365	CD4_MOUSE	P97792 mus musculu
8	124	8.4	306	CD80_MOUSE	Q00609 mus musculu
9	120.5	8.2	403	CD33_MOUSE	Q63994 mus musculu
10	119.5	8.1	322	ICOL_MOUSE	O99138 mus musculu
11	118.5	8.1	551	SRB1_HUMAN	O15389 homo sapien
12	116	7.9	459	CD4_MOUSE	P46630 oryctolagus
13	114	7.8	246	CD4_MOUSE	O61885 mus musculu
14	112.5	7.7	463	SRB1_HUMAN	O99138 mus musculu
15	109.5	7.5	457	CD4_MOUSE	O28037 saimiri sci
16	108.5	7.4	245	CD4_MOUSE	O63345 rattus norv
17	105	7.2	215	CD4_MOUSE	P54900 rattus norv
18	105	7.2	354	CD4_MOUSE	P03994 rattus norv
19	104	7.1	354	CD4_MOUSE	P10915 homo sapien
20	104	7.1	513	SRB1_MOUSE	P97797 m protein-t
21	103.5	7.1	354	CD4_MOUSE	O28381 equus caball
22	102.5	7.0	356	CD4_MOUSE	O99138 mus musculu
23	102.5	7.0	457	CD4_MOUSE	P05540 rattus norv
24	102	6.9	457	CD4_MOUSE	P05332 mus musculu
25	101.5	6.9	458	CD4_MOUSE	P01730 homo sapien
26	100.5	6.8	458	CD4_MOUSE	O08338 cercoptithec
27	100.5	6.8	458	CD4_MOUSE	P71604 pan troglod
28	100.5	6.8	458	CD4_MOUSE	O46631 bos taurus
29	100.5	6.8	506	CD4_MOUSE	P20138 homo sapien
30	100	6.8	364	CD4_MOUSE	O96137 homo sapien
31	100	6.8	458	CD4_MOUSE	P79184 macaca fasc
32	99.5	6.8	458	CD4_MOUSE	P16003 macaca mula
33	99.5	6.8	458	CD4_MOUSE	

34	99.5	6.8	458	CD4_MOUSE	Q08340 macaca neme
35	99	6.7	215	CD4_MOUSE	O60939 homo sapien
36	98.5	6.7	345	CD4_MOUSE	P98172 homo sapien
37	98.5	6.7	354	CD4_MOUSE	P55252 bos taurus
38	98.5	6.7	354	CD4_MOUSE	P10859 sus scrofa
39	98	6.7	575	CD4_MOUSE	O99138 mus musculu
40	98	6.7	597	CD4_MOUSE	O99138 mus musculu
41	97	6.6	686	CD4_MOUSE	O99138 mus musculu
42	97	6.6	1914	CD4_MOUSE	O15746 homo sapien
43	96.5	6.6	1005	CD4_MOUSE	P29853 aspergillus
44	96	6.5	345	CD4_MOUSE	P52795 mus musculu
45	96	6.5	569	CD4_MOUSE	Q92093 mus musculu

## ALIGNMENTS

RESULT 1  
A33\_HUMAN STANDARD; PRT; 319 AA.  
AC Q99795;  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Cell surface A33 antigen precursor (Glycoprotein A33).  
GN GPR33.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Colon carcinoma;  
RX MEDLINE=97165045; PubMed=9012807;  
RA Heath J.K., White S.J., Johnstone C.N., Catimel B., Simpson R.J.,  
RA Moritz R.L., Tu G.-F., Ji H., Whitehead R.H., Groenen L.C.,  
RA Scott A.M., Ritter G., Cohen L., Old L.J., Nice E.C.,  
RA Burgess A.W.;  
RT "The human A33 antigen is a transmembrane glycoprotein and a novel  
RT member of the immunoglobulin superfamily.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:469-474(1997).  
RN [2]  
RP POST-TRANSLATIONAL MODIFICATIONS.  
RX MEDLINE=97396159; PubMed=9245713;  
RA Ritter G., Cohen L.S., Nice E.C., Catimel B., Burgess A.W.,  
RA Moritz R.L., Ji H., Heath J.K., White S.J., Old L.J.,  
RA Simpson R.J.;  
RT "Characterization of posttranslational modifications of human A33  
RT antigen, a novel palmitoylated surface glycoprotein of human  
RT Gastrointestinal epithelium.";  
RL Biochem. Biophys. Res. Commun. 236:682-686(1997).  
CC -1- FUNCTION: MAY PLAY A ROLE IN CELL-CELL RECOGNITION AND SIGNALING.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN NORMAL GASTROINTESTINAL  
CC EPITHELIUM AND IN 95% OF COLON CANCERS.  
CC -1- PTM: N-GLYCOSYLATED, CONTAINS APPROXIMATELY 8 KDA OF N-LINKED  
CC CARBOHYDRATE.  
CC -1- PTR: PALMITOYLATED.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL, U79725; AAC50957.1; -;  
CC Genbank, HGNC:4445; GPR33.  
CC MIM, 602171; -;  
CC GO, GO:0005888; C:proteoglycan integral to plasma membrane; TAS.

DR GO:0004872; F:receptor activity; TAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS00835; IG LIKE; 2.  
 KW Immunoglobulin domain, lipoprotein, Palmitate; Glycoprotein;  
 KW Transmembrane; Signal; Antigen.  
 FT SIGNAL 1 21  
 FT CHAIN 1 21  
 FT FT 22 319 CELL SURFACE A33 ANTIGEN.  
 FT DOMAIN 22 235 EXTRACELLULAR (POTENTIAL).  
 FT TRAMMEM 236 256 POTENTIAL.  
 FT DOMAIN 257 319 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 22 134 IG-LIKE V-TYPE.  
 FT DOMAIN 140 227 IG-LIKE C2-TYPE.  
 FT DOMAIN 258 261 POLY-CYS.  
 FT DISULFID 43 117 POTENTIAL.  
 FT DISULFID 146 222 POTENTIAL.  
 FT DISULFID 162 211 POTENTIAL.  
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 319 AA; 35632 MW; 9BFC7AFA5C2408E CRC64;  
 Query Match 10.7%; Score 157; DB 1; Length 319;  
 Beest Local Similarity 22.7%; Pred. No. 5.8e-06;  
 Matches 73; Conservative 40; Mismatches 116; Indels 92; Gaps 12;  
 QY 6 TLNCLVLLQLLLANSLDEGVKNAVLPSCSYTLPTSGTLVPMCGKGFPMSCQTN 65  
 DB 11 TLCAVAVTDAISVERPPQVLRASQSKSVTLPTCTHTSTSSR-----EGLIQMDKL-- 61  
 QY 66 ELRLDERNNVYQKSRVQLKQDLNK-----GVSLLIKNVTLDDHGYCCRIQ 114  
 DB 62 -LLTTERVAVIWPFSKNVYHGLYKNRYSINNAEQSDASITIDLTWADNGYECV- 119  
 QY 115 PFGILNND-----KKLEKLDIKAVKTP----- 137  
 DB 120 --SLMSDELGNTKSRVRLVLPSPKPGEGIGETITGNITQTLTCKSGKSPQYSWKR 177  
 QY 138 -----AQTAGDSTTASPRTLTTERNGSEFTGLTLNHNNGTKISTWADEIKDSG 187  
 DB 178 VNIINQEPRLAQPASQ--PVLKINISTDTSG--YYITSSNEBETQCINTAVRSPS 232  
 QY 188 ETRPAIHIGVGSAGTLALIIIGVILKWSG-----KKKLSISLITLANLPFG 239  
 DB 233 NNVALYVGIAVGVAAL---IIIGITII---YCCCCRGKDDNTEDKEDARPNREAYEPE 286  
 QY 240 GLANNAVIRSEENITYTEE 260  
 DB 287 QLRRLS--REREEDDYROE 305  
 RESULT 2  
 CXAR\_HUMAN STANDARD; PRT; 365 AA.  
 ID CXAR\_HUMAN STANDARD; PRT; 365 AA.  
 AC P78310; O00694;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Coxsackievirus and adenovirus receptor precursor (Coxsackievirus B-  
 adenovirus receptor) (hCAR) (CVB3 binding protein).  
 GN CXAR OR CAR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID:9606;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=97190109; PubMed=9036860;  
 RA Bergelson J.M., Cunningham J.A., Droguett G., Kurt-Jones E.,  
 RA Krithivas A., Hong J.S., Horwitz M.S., Crowell R.L., Finberg R.W.;

RT "Isolation of a common receptor for Coxsackie B viruses and  
 RT adenoviruses 2 and 5.";  
 RT Science 275:1320-1323 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE=97250541; PubMed=9066397;  
 RA Tomko R.P., Xu R., Philipson L.;  
 RT "hCAR and hCAR: the human and mouse cellular receptors for subgroup C  
 RT adenoviruses and group B coxsackieviruses.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 94:3352-3356 (1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE=20008750; PubMed=10543405;  
 RA Bowles K.R., Gibson J., Wu J., Shaffer L.G., Towbin J.A.,  
 RA Bowles N.E.;  
 RT "Genomic organization and chromosomal localization of the human  
 RT Coxsackievirus B-adenovirus receptor gene.";  
 RT Hum. Genet. 105:354-359 (1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Anderson C.W., Kieleczawa J., Dunn J.J., Freimuth P.;  
 RT "Sequence and expression of CXAR, the human gene for the  
 RT coxsackievirus and adenovirus receptor.";  
 RT Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Anderson B., Tomko R., Anderson K., Darban H., Oncu D., Mizra M.,  
 RA Solterdrant K., Sonhammer E., Philipson L.;  
 RT "Putative regulatory domains in the human and mouse CAR genes.";  
 RT Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cervix;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemmer C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Burow K.H., Scheffer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares W.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalon D.K., Wuzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- FUNCTION: SERVES AS A RECEPTOR FOR GROUP B COXSACKIEVIRUSES AND  
 CC SUBGROUP C OF ADENOVIRUSES (AD2 AND AD5).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; Y07593; CAA68868.1; -;  
 DR EMBL; U90716; AAC51234.1; -;  
 DR EMBL; AF169366; AAF05908.1; -;  
 DR EMBL; AF169360; AAF05908.1; JOINED.  
 DR EMBL; AF169361; AAF05908.1; JOINED.  
 DR EMBL; AF169362; AAF05908.1; JOINED.



DR EMBL; AF169363; AAF05908.1; JOINED.  
 DR EMBL; AF169364; AAF05908.1; JOINED.  
 DR EMBL; AF169365; AAF05908.1; JOINED.  
 DR EMBL; AF200465; AAF24344.1; -  
 DR EMBL; AF242865; AAG01088.1; -  
 DR EMBL; AF242862; AAG01088.1; JOINED.  
 DR EMBL; AF242864; AAG01088.1; JOINED.  
 DR EMBL; BC003684; AAH03684.1; -  
 DR EMBL; BC010536; AAH10536.1; -  
 DR PDB; 1FAJ; 13-JUL-01.  
 DR PDB; 1FSW; 08-NOV-00.  
 DR PDB; 1KAC; 24-NOV-99.  
 DR Genew; HGNC:2559; CXADR.  
 DR MIM; 602621; -  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0004872; F:receptor activity; TAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00408; Igc2; 1.  
 DR PROSITE; PSS0835; IG\_LIKE; 2.  
 KW Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; Signal;  
 KW Repeat; 3d-structure.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 365 COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR.  
 FT DOMAIN 20 237 EXPACELLULAR (POTENTIAL).  
 FT TRANSMEM 238 258 POTENTIAL.  
 FT DOMAIN 259 365 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 20 134 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 141 228 IG-LIKE C2-TYPE 2.  
 FT DISULFID 41 120 BY SIMILARITY.  
 FT DISULFID 162 212 BY SIMILARITY.  
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 365 AA; 40029 MW; AB01C6346CB7EE64 CRC64;  
 Query March 10.0%; Score 147.5; DB 1; Length 365;  
 Best Local Similarity 23.5%; Pred. No. 4.3e-05;  
 Matches 63; Conservative 47; Mismatches 101; Indels 57; Gaps 11;  
 QY 7 LNCVLLLLQLL-LARSL-----EDGYKVEGKNAFLPSCYTLPTSGTLVPMCWKGCPW 60  
 DB 4 LLCFVLGCVVDFARSLITTPREMIKAKGETAYLPCKFTLS-----PEDQGPLDIEM 57  
 QY 61 SGCNELLRDERVATYQKSSRY-----QLKG-----DLNKDVSIIINVTLDHGT 108  
 DB 58 LISPADQKQDVIIISGDKIYDDYYPDLKGRVHFTSNLKSQDASINVTNLQISDIGN 117  
 QY 109 YCCRI-QPFGIMNDKXLELKIDKAA-----KVTPAQTA-----HGDS 145  
 DB 118 YQCVKAKAPGVA-N-KKHLVVLVPSGARCYVDSSEISGSPFKICEPKSSLPLOYEWQ 176  
 QY 146 TTASPRFLTERNGSETQTLVTLNNNGTKSTWADEIKDSGETIRTAHIGV---GVSA 202  
 DB 177 KLSQSQKMPSTWLAEMTSSVSVNASESEYGTSCYRANRVSDDCLRLNVVPPSKA 236  
 QY 203 GLTLLIIGVLI-----LKKYSCKKK 224  
 DB 237 GLIAGAIITLALALGLIIFCCRRK 264  
 RESULT 3  
 SRBI HUMAN STANDARD; PRT; 398 AA.  
 AC 000241; O8TB12; Q9H1U5; Q9Y4V0;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Signal-regulatory protein beta-1 precursor (SRP-beta-1).  
 GN SRPB1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Placenta;  
 RX MEDLINE=97215901; PubMed=9062191;  
 RA Kharitonkov A., Chen Z., Sures I., Wang H., Schilling J.,  
 RA Ullrich A.;  
 RT "A family of proteins that inhibit signalling through tyrosine kinase  
 RT receptors";  
 RL Nature 386:181-186(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor P.D., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leharasliho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McKerray A.A.,  
 RA Malne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Philimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
 RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.W.,  
 RA Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Klausner R.L., Peingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heileh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stepieton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uscin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP INTERACTION WITH TYROBP AND SYK.  
 RX PubMed=11169422;  
 RA Tomasello E., Cant C., Buehring H.-J., Vely F., Andre P., Seifert M.,  
 RA Ullrich A., Vinyer E.;  
 RT "Association of signal-regulatory proteins beta with KAPAP/DAP-12.";  
 RL Eur. J. Immunol. 30:2147-2156(2000).  
 RN [5]

Query	Match	Score	DB 1	Length	DB 2	Score	DB 3	Length	DB 4	Score	DB 5	Length	DB 6	Score	DB 7	Length	DB 8	Score	DB 9	Length	DB 10	Score	DB 11	Length	DB 12	Score	DB 13	Length	DB 14	Score	DB 15	Length	DB 16	Score	DB 17	Length	DB 18	Score	DB 19	Length	DB 20	Score	DB 21	Length	DB 22	Score	DB 23	Length	DB 24	Score	DB 25	Length	DB 26	Score	DB 27	Length	DB 28	Score	DB 29	Length	DB 30	Score	DB 31	Length	DB 32	Score	DB 33	Length	DB 34	Score	DB 35	Length	DB 36	Score	DB 37	Length	DB 38	Score	DB 39	Length	DB 40	Score	DB 41	Length	DB 42	Score	DB 43	Length	DB 44	Score	DB 45	Length	DB 46	Score	DB 47	Length	DB 48	Score	DB 49	Length	DB 50	Score	DB 51	Length	DB 52	Score	DB 53	Length	DB 54	Score	DB 55	Length	DB 56	Score	DB 57	Length	DB 58	Score	DB 59	Length	DB 60	Score	DB 61	Length	DB 62	Score	DB 63	Length	DB 64	Score	DB 65	Length	DB 66	Score	DB 67	Length	DB 68	Score	DB 69	Length	DB 70	Score	DB 71	Length	DB 72	Score	DB 73	Length	DB 74	Score	DB 75	Length	DB 76	Score	DB 77	Length	DB 78	Score	DB 79	Length	DB 80	Score	DB 81	Length	DB 82	Score	DB 83	Length	DB 84	Score	DB 85	Length	DB 86	Score	DB 87	Length	DB 88	Score	DB 89	Length	DB 90	Score	DB 91	Length	DB 92	Score	DB 93	Length	DB 94	Score	DB 95	Length	DB 96	Score	DB 97	Length	DB 98	Score	DB 99	Length	DB 100	Score	DB 101	Length	DB 102	Score	DB 103	Length	DB 104	Score	DB 105	Length	DB 106	Score	DB 107	Length	DB 108	Score	DB 109	Length	DB 110	Score	DB 111	Length	DB 112	Score	DB 113	Length	DB 114	Score	DB 115	Length	DB 116	Score	DB 117	Length	DB 118	Score	DB 119	Length	DB 120	Score	DB 121	Length	DB 122	Score	DB 123	Length	DB 124	Score	DB 125	Length	DB 126	Score	DB 127	Length	DB 128	Score	DB 129	Length	DB 130	Score	DB 131	Length	DB 132	Score	DB 133	Length	DB 134	Score	DB 135	Length	DB 136	Score	DB 137	Length	DB 138	Score	DB 139	Length	DB 140	Score	DB 141	Length	DB 142	Score	DB 143	Length	DB 144	Score	DB 145	Length	DB 146	Score	DB 147	Length	DB 148	Score	DB 149	Length	DB 150	Score	DB 151	Length	DB 152	Score	DB 153	Length	DB 154	Score	DB 155	Length	DB 156	Score	DB 157	Length	DB 158	Score	DB 159	Length	DB 160	Score	DB 161	Length	DB 162	Score	DB 163	Length	DB 164	Score	DB 165	Length	DB 166	Score	DB 167	Length	DB 168	Score	DB 169	Length	DB 170	Score	DB 171	Length	DB 172	Score	DB 173	Length	DB 174	Score	DB 175	Length	DB 176	Score	DB 177	Length	DB 178	Score	DB 179	Length	DB 180	Score	DB 181	Length	DB 182	Score	DB 183	Length	DB 184	Score	DB 185	Length	DB 186	Score	DB 187	Length	DB 188	Score	DB 189	Length	DB 190	Score	DB 191	Length	DB 192	Score	DB 193	Length	DB 194	Score	DB 195	Length	DB 196	Score	DB 197	Length	DB 198	Score	DB 199	Length	DB 200	Score	DB 201	Length	DB 202	Score	DB 203	Length	DB 204	Score	DB 205	Length	DB 206	Score	DB 207	Length	DB 208	Score	DB 209	Length	DB 210	Score	DB 211	Length	DB 212	Score	DB 213	Length	DB 214	Score	DB 215	Length	DB 216	Score	DB 217	Length	DB 218	Score	DB 219	Length	DB 220	Score	DB 221	Length	DB 222	Score	DB 223	Length	DB 224	Score	DB 225	Length	DB 226	Score	DB 227	Length	DB 228	Score	DB 229	Length	DB 230	Score	DB 231	Length	DB 232	Score	DB 233	Length	DB 234	Score	DB 235	Length	DB 236	Score	DB 237	Length	DB 238	Score	DB 239	Length	DB 240	Score	DB 241	Length	DB 242	Score	DB 243	Length	DB 244	Score	DB 245	Length	DB 246	Score	DB 247	Length	DB 248	Score	DB 249	Length	DB 250	Score	DB 251	Length	DB 252	Score	DB 253	Length	DB 254	Score	DB 255	Length	DB 256	Score	DB 257	Length	DB 258	Score	DB 259	Length	DB 260	Score	DB 261	Length</
-------	-------	-------	------	--------	------	-------	------	--------	------	-------	------	--------	------	-------	------	--------	------	-------	------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	----------

[illegible]

RP ARG-54; ALA-57; GLY-75; ASP-95; LEU-96; ASN-100; ARG-107; GLY-109 AND  
 RP VAL-131.  
 RC TISSUE=Brain;  
 RX MEDLINE=20053880; PubMed=10585853;  
 RA Sano S.-I., Ohnishi H., Kubota M.;  
 RT "Gene structure of mouse B17/SHPS-1.";  
 RL Biochem. J. 344:667-675(1999).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stevrides G., Almeida J.P., Babbage A.K., Baggaley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cordy N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.T., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lehar V.L., Martin S.T., McCormachie L.J., McIay K., Lovell J.D.,  
 RA Marsh V.L., Martin S.T., McCormachie L.J., McIay K., Lovell J.D.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramay H.,  
 RA Rice C.M., Ross M.T., Scott C.B., Senta H.R., Showkhen R., Sims S.,  
 RA Skuse C.D., Smith W.L., Soderlund C., Stewart C.A., Sultson J.E.,  
 RA Swann R.M., Symcote N., Taylor R.M., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A.C., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS LEU-44; THR-50;  
 RP THR-53; ARG-54; ALA-57; GLY-75; ASP-95; LEU-96; ASN-100; ARG-107;  
 RP GLY-109 AND VAL-131.  
 RC TISSUE=Brain, Kidney, and Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [7]  
 RP FUNCTION AND INTERACTIONS WITH FYB, SCAP2 AND PTK2B.  
 RX MEDLINE=99401000; PubMed=10469599;  
 RA Tims J.F., Swanson K.D., Marie-Cardine A., Raab M., Rudd C.E.,  
 RA Schraven B., Neel B.G.;  
 RT "SHPS-1 is a scaffold for assembling distinct adhesion-regulated  
 RT multi-protein complexes in macrophages.";  
 RL Curr. Biol. 9:927-930(1999).  
 RN [8]  
 RP PHOSPHORYLATION BY JAK2, AND INTERACTIONS WITH PRPN1 AND JAK2.

RX MEDLINE=20428742; PubMed=10842184;  
 RA Stofega M.R., Argersinger L.S., Wang H., Ulrich A., Carter-Su C.;  
 RT "Negative regulation of growth hormone receptor/JAK2 signaling by  
 RT signal regulatory protein alpha.";  
 RL J. Biol. Chem. 275:28222-28229(2000).  
 RN [9]  
 RP FUNCTION AND INTERACTION WITH CD47.  
 RX MEDLINE=21400825; PubMed=11509594;  
 RA Latour S., Tanaka H., Demeure C., Mateo V., Rubio M., Brown E.J.,  
 RA Maliszewski G., Lindberg F.P., Oldenborg A., Ulrich A.,  
 RA Delespere G., Sarrail M.;  
 RT "Bidirectional negative regulation of human T and dendritic cells by  
 RT CD47 and its cognate receptor signal-regulator protein-alpha:  
 RT down-regulation of IL-12 responsiveness and inhibition of dendritic  
 cell activation.";  
 RL J. Immunol. 167:2547-2554(2001).  
 CC -1- FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Acts  
 CC as docking protein and induces translocation of PRPN6, PRPN1 and  
 CC other binding partners from the cytosol to the plasma membrane.  
 CC Supports adhesion of cerebellar neurons, neurite outgrowth and  
 CC glial cell attachment. May play a key role in intracellular  
 CC signaling during synaptogenesis and in synaptic function (by  
 CC exocytosis). Involved in the negative regulation of receptor  
 CC tyrosine kinase-coupled cellular responses induced by cell  
 CC adhesion, growth factors or insulin. Mediates negative regulation  
 CC of phagocytosis, mast cell activation and dendritic cell  
 CC activation. CD47 binding prevents maturation of immature dendritic  
 CC cells and inhibits cytokine production by mature dendritic cells.  
 CC -1- SUBUNIT: Binds PRPN1 when tyrosine-phosphorylated, except in  
 CC microtubules, where it primarily binds PRPN6. Binds GRB2 in vitro.  
 CC Binds FcR (by similarity). Binds JAK2 irrespective of its  
 CC phosphorylation status and forms a stable complex. Binds SCAP1  
 CC and/or SCAP2. The resulting complex recruits FYB. Binds PTK2B.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1;  
 CC IsoId=P78324-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=P78324-2; Sequence=VSP\_007030;  
 CC Note=No experimental confirmation available;  
 CC Name=3;  
 CC IsoId=P78324-3; Sequence=VSP\_007029;  
 CC Note=No experimental confirmation available;  
 CC -1- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in brain.  
 CC Detected on myeloid cells, but not T cells. Detected at lower  
 CC levels in heart, placenta, lung, testis, ovary, colon, liver,  
 CC small intestine, prostate, spleen, kidney, skeletal muscle and  
 CC pancreas.  
 CC -1- PTM: N-glycosylated.  
 CC -1- PTM: Phosphorylated on tyrosine residues in response to  
 CC stimulation with BGF, growth hormone, insulin and PDGF.  
 CC Dephosphorylated by PRPN1.  
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; D86043; BAI12974.1; -;  
 CC EMBL; Y10375; CAA71403.1; -;  
 CC EMBL; Y1047; CAA71944.1; ALT INIT.  
 CC EMBL; AB023430; BAA87929.1; -;  
 CC EMBL; AC004832; ARI19260.1; -;  
 CC EMBL; AL034562; CAB38874.1; -;  
 CC EMBL; AL049634; CAB46662.1; ALT-SEQ.  
 CC EMBL; AL117335; CAC12723.1; -;  
 CC EMBL; BC026692; AAB26692.1; -;

```

DR EMBL: BC033092; AAH33092.1; -
DR EMBL: BC038510; AAH38510.1; -
DR PIR: JCS287; JCS287.
DR Genew: HGNC:9662; PRPN51.
DR MIM: 602461; -
DR GO: 0005886; Cytoplasm membrane; TAS.

Query Match      8.9%; Score 131; DB 1; Length 501;
Best Local Similarity 22.5%; Pred. No. 0.0016;
Matches 67; Conservative 45; Mismatches 124; Indels 62; Gaps 13;

QY 13 LLOLLARS- - - - - EDGKVEGVGNAYLPCSYT--LPTSGTLVPMCWG 54
DB 14 LLLCLLAASCAAGVAGEELVIOIPDKSVSAAGESALHCTVSLIVG- - - - - PLOMF 69
QY 55 KQPCWMSQCTNELRLTDE- - - - - RNVYQSSRRYQKGLNKGDVSLIKVTLDDHGT 109
DB 70 RGAGP- - - - - ARELIYQKXGHPFRVTVTVESTKRE- - - - - NMDFISISNITPADAGTY 119
QY 110 CCRIOFPGMLNDKLE- - - - - LKLDIKAAVTPAQTHGSDTASPR- - - - - TLTERNGSET 162
DB 120 YC-VKFRKGSPTPEFGSGAGTSLVAKPSAP- - - - - VVSGPAARATPQHTVSFTCESHGFSP 176
QY 163 QTLVTLNNHNTKISTWMADEIKDSCGRTTAIHIGVAGAGTLAILIIGVLIKMYSCCK 222
DB 177 RDTLTKMFNGNELSDFOINVDVPGESVSYSIHSTAKV- - - - - VLTREDVHS 223
QY 223 KKLSSLSLITLANLPFGGLAAGAVRISEENITYTEENVYEVENSNEYCYVNSQOP 280
DB 224 QVICVAHVTLQGDPLRGTAANLSET-IRVPPTLEVTVQGV-BAENQVNTVCVRKPYF 279

RESULT 5
SRB2_HUMAN STANDARD; PRT; 387 AA.
AC Q9P1W8; Q8WMA5; Q9NOK8;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Signal-regulatory protein beta-2 precursor (SIRP-beta-2) (SIRP-b2).
GN SIRPB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RX SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Placenta;
RA MEDLINE=21036165; PubMed=1185750;
RA Ichigotani Y., Matsuda S., Machida K., Oshima K., Iwamoto T.,
RA Yamaki K., Hayakawa T., Hamaguchi M.;
RT "Molecular cloning of a novel human gene (SIRP-B2) which encodes a new
RT member of the SIRP/SHPS-1 protein family.";
RT J. Hum. Genet. 45:378-382(2000).
[2]
RX SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RX Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RX Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RX Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RX Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RX Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RX Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RX Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RX Coulson A.G., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RX Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RX Graham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
RX Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RX Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RX Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RX Levaslahti M.H., Leversta M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RX Marsh V.L., Martin S.L., McCormack L.J., Mcley K., McMurray A.A.,
RX Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

```

```

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,
RA Rice C.M., Ross M.T., Scott C.E., Senra H.K., Showkhen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sultson J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thoype A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilting L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RT Nature 414:865-871(2001).
[3]
RX SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RX Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RX Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RX Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RX Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RX Diachenko L., Marinova K., Farmer A.A., Rubin G.M., Hong L.,
RX Stabileton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RX Brownstein M.J., Usdin T.B., Toshiyuki S., Carinici P., Prange C.,
RX Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RX Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RX Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RX Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Probable immunoglobulin-like cell surface receptor.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9P1W8-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9P1W8-2; Sequence=VSP_007027;
CC Name=3;
CC IsoId=Q9P1W8-3; Sequence=VSP_007028;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Detected in liver, and at very low levels in
CC brain, heart, lung, pancreas, kidney, placenta and skeletal
CC muscle.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C1-type domains.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licences@isb-sib.ch).
DR EMBL: AB042624; BAA95692.1; -
DR EMBL: AL038804; CAC00474.1; -
DR EMBL: BC020629; AAH20629.1; ALT_INIT.
DR Genew: HGNC:15757; SIRPB2.
DR MIM: 605466; -
DR GO: 0007267; P:cell-cell signaling; TAS.
DR GO: 0007242; P:intracellular signaling cascade; TAS.
DR GO: 0008285; P:negative regulation of cell proliferation; TAS.
DR InterPro: IPR007110; IG-1-like.
DR InterPro: IPR003597; IG_c1.
DR InterPro: IPR003006; IG_MHC.
DR Pfam: PF00047; Ig_3.

```

DR SMART; SM00407; IGc1; 2.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; FALSE NEG.  
 KM Repeat; signal; Transmembrane; Immunoglobulin domain; Glycoprotein;  
 KM Alternative splicing.  
 FT SIGNAL 1 28  
 FT CHAIN 29 387  
 FT DOMAIN 29 360  
 FT TRANSMEM 361 383  
 FT DOMAIN 384 387  
 FT DOMAIN 29 137  
 FT DOMAIN 146 245  
 FT DOMAIN 252 340  
 FT DISULFID 53 119  
 FT DISULFID 168 226  
 FT CARBOHYD 271 329  
 FT CARBOHYD 243 243  
 FT CARBOHYD 268 268  
 FT CARBOHYD 309 309  
 FT CARBOHYD 317 317  
 FT VARSPIC 1 33  
 FT VARSPIC 144 360  
 FT CONFLICT 263 263  
 FT CONFLICT 286 286  
 SQ SEQUENCE 387 AA; 42495 MW; F7F20C9F86B0E64B CRC64;  
 Query Match 8.8%; Score 129; DB 1; Length 387;  
 Best Local Similarity 23.1%; Pred. No. 0.0017;  
 Matches 65; Conservative 38; Mismatches 90; Indels 88; Gaps 13;  
 QY 12 LLLDLLARSLDEGK-----VEGKAYLP CST--LPISGTVPMCKGKF 57  
 DB 15 LLLTLGLTEVAGEEELQMIQPEKLLLVYGTATATLCTVTSLLPVG---PVLMFEGV 70  
 QY 58 CPMSQCTNELLRTBERNTVYOKSS---RYOLKEDL---NKGVSLLIKNYTLDDHGYTC 110  
 DB 71 GP-----GRELIYNQKXGHPRTVTSVDTLRNNMDSIRISITPADVGYTY 118  
 QY 111 C-----RIQF---PGLMNDKLEIKDIDK-----AAKVTPAQTAGDSTTAS 149  
 DB 119 CVKFKKSPENVEFGSPG-----TEWALGAKPSAPVGLPAARTPEHTV----- 164  
 QY 150 PRITLTERNGSEQTQVTLVTHNNGTKISTWADETKQSETRITRAIHIGVGSAGITLALI 209  
 DB 165 --SFTCSHGSPSPDITLKWFKNENELSDFTQNVDPGQSVAYSIRSTARV----- 213  
 QY 210 IGVLLIKWYSCKKKLSLITLANLPGGIAN-AGAVRI 249  
 DB 214 ----VLDPWDVRSQVICEVAHTVLOQDPLRGTANLSEAIRV 250  
 RESULT 6  
 SIL8\_HUMAN STANDARD; PRT; 499 AA.  
 ID AC Q9NYZ4;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Sialic acid binding Ig-like lectin 8 precursor (Siglec-8)  
 DE (Sialoadhesin family member-2) (SAF-2).  
 GN SIGLEC8 OR SAF2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=20549027; PubMed=11095983;  
 RA Fousheas G., Yousef G.M., Diamandis E.P.;  
 RT "Molecular characterization of a siglec8 variant containing

RT Cytoplasmic tyrosine-based motifs, and mapping of the siglec8 gene.";  
 RL Biochem. Biophys. Res. Commun. 278:775-781(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RX MEDLINE=20314554; PubMed=10856141;  
 RA Kikly K.K., Bochner B.S., Freeman S.D., Tan K.B., Gallagher K.T.,  
 RA D'Alessio K.J., Holmes S.D., Abrahamson J.A., Erickson-Willer C.L.,  
 RA Murdoch P.R., Tachimoto H., Schleimer R.P., White J.R.;  
 RT "Identification of SAF-2, a novel siglec expressed on eosinophils,  
 RT mast cells, and basophils.";  
 RL J. Allergy Clin. Immunol. 105:1093-1100(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RX TISSUE=Eosinophil;  
 RC MEDLINE=20092847; PubMed=10625619;  
 RA Floyd H., Ni J., Cornish A.L., Zeng Z., Liu D., Carter K.C., Steel J.,  
 RA Crocker P.R.;  
 RT "Siglec-8, A novel eosinophil-specific member of the immunoglobulin  
 RT superfamily.";  
 RL J. Biol. Chem. 275:861-866(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RA Floyd H., Zhang J.Q., Crocker P.R.;  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Putative adhesion molecule that mediates sialic-acid  
 CC dependent binding to cells. Preferentially binds to alpha2,3-  
 CC linked sialic acid. Also binds to alpha2,6-linked sialic acid. The  
 CC sialic acid recognition site may be masked by cis interactions  
 CC with sialic acids on the same cell surface.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Name=1; Synonyms=long;  
 CC IsoId=Q9NYZ4-1; Sequence=displayed;  
 CC Name=2;  
 CC IsoId=Q9NYZ4-2; Sequence=VSP\_002559;  
 CC Name=3;  
 CC IsoId=Q9NYZ4-3; Sequence=VSP\_002560;  
 CC -1- TISSUE SPECIFICITY: Expressed specifically on eosinophils.  
 CC -1- DOMAIN: Contains an intracytoplasmic motif referred as  
 CC immunoreceptor tyrosine-based inhibitor motif (ITIM). This motif  
 CC is involved in downmodulation of cellular functions as the  
 CC termination of the immune response.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
 CC (SIALIC ACID BINDING IG-LIKE LECTIN) SUBFAMILY.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC EMBL; AF287892; AAC00573.1; -  
 DR EMBL; AF223403; AAF34702.1; -  
 DR EMBL; AF195092; AAF27622.1; -  
 DR EMBL; AF310234; AAK55140.1; -  
 DR Genew; HGNC:10877; SIGLEC8.  
 DR MIM; 605639; -  
 DR GO; GO:0016021; C:integral to membrane; TAS.  
 DR GO; GO:0005530; F:lectin; TAS.  
 DR GO; GO:0004888; F:transmembrane receptor activity; TAS.  
 DR GO; GO:0007165; F:signal transduction; TAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig\_2.  
 DR SMART; SM00408; IGc2; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 2.  
 KW Cell adhesion; lectin; Transmembrane; signal; Glycoprotein;



```

Db      146 CFVDSGEIGNDFLXKCEPKSGSLPQFEMQKSDSQMPRLAEMTSPVSYKNA5SE 205
Qy      157 RNSGETQTLVTHNNNGTKISTMADEIKDSEGTIRTHIHG-VGVSAGLTLAI-IGVL 213
Db      206 YSGTYS---CTVQNRVG-----SDQCMRLDVVPSPSRAGTIAGAVIGTLALVLIGAI 256
Qy      214 ILKWSCKKKK 224
Db      257 L---FCCHRRK 264

RESULT 8
CD80_MOUSE STANDARD; PRT; 306 AA.
AC 000609;
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE T lymphocyte activation antigen CD80 precursor (activation B7-1
  antigen) (B7).
GN CD80 OR B7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=B-cell;
RX MEDLINE=91341422; PubMed=1714935;
RA Gray G.S., Freeman G.J., Gimmi C.D., Lombard D.B., Zhou L.J.,
RA White M., Fingerhuth J.D., Gribben J.G., Nadler L.M.;
RT "Structure, expression, and T cell costimulatory activity of the
  murine homologue of the human B lymphocyte activation antigen B7.",
  J. Exp. Med. 174:625-631 (1991).
RL [2]
RN SEQUENCE FROM N.A.
RC TISSUE=B-cell;
RX MEDLINE=93307789; PubMed=7686531;
RA Selvakumar A., White P.C., Dupont B.;
RT "Genomic organization of the mouse B-lymphocyte activation antigen
  B7.",
  Immunogenetics 38:292-295 (1993).
RL [3]
RN SEQUENCE FROM N.A.
RC TISSUE=B-cell;
RX MEDLINE=93307789; PubMed=7686531;
RA Selvakumar A., White P.C., Dupont B.;
RT "Genomic organization of the mouse B-lymphocyte activation antigen
  B7.",
  Immunogenetics 38:292-295 (1993).
CC -1- FUNCTION: INVOLVED IN THE COSTIMULATORY SIGNAL ESSENTIAL FOR T
  LYMPHOCYTES ACTIVATION. T CELL PROLIFERATION AND CYTOKINE
  PRODUCTION IS INDUCED BY THE BINDING OF CD28 OR CTLA-4 TO THIS
  RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED ON ACTIVATED B CELLS, GAMMA
  INTERFERON STIMULATED MONOCYTES AND NONCIRCULATING B-CELL
  MALIGNANCIES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED BETWEEN 4 AND 12 HOURS POST-
  ACTIVATION. PROTEIN WAS DETECTED AT CELL SURFACE AT 24 HOURS AND
  IT'S EXPRESSION WAS MAXIMAL FROM 48 TO 72 HOURS POST-ACTIVATION.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X60958; CA443291.1; -
DR EMBL, L12589; AAA37240.1; ALT SEQ.
DR EMBL, L12585; AAA37240.1; JOINED.
DR EMBL, L12586; AAA37240.1; JOINED.
DR EMBL, L12587; AAA37240.1; JOINED.
DR EMBL, L12588; AAA37240.1; JOINED.
DR PIR, I49503; I49503.
DR HSSP, P33681; IDR9.

```

```

DR MGD; MG1:101775; C880.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; Ig; 1.
DR PROSITE; PS50835; Ig_LIKE; 2.
KW Immunoglobulin domain; T-cell; Glycoprotein; Signal; Transmembrane;
  Receptor.
FT SIGNAL 1 37
FT CHAIN 38 306 T LYMPHOCYTE ACTIVATION ANTIGEN CD80.
FT DOMAIN 38 246 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 247 268 POTENTIAL.
FT DOMAIN 269 306 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 38 135 IG-LIKE V-TYPE.
FT DOMAIN 148 229 IG-LIKE C2-TYPE.
FT DOMAIN 227 246 IG-HINGE LIKE (POTENTIAL).
FT DISULFID 54 119 POTENTIAL.
FT DISULFID 165 219 POTENTIAL.
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 306 AA; 34589 MW; 1DBADE0931B84C62 CRC64;

Query Match 8.44; Score 124; DB 1; Length 306;
Best Local Similarity 24.08; Pred. No. 0.0033;
Matches 67; Conservative 32; Mismatches 90; Indels 90; Gaps 12;

Qy 11 LLLQLLLAR-----SLGQYKVEGKNAYLPSCSYTLPTSG-----TLVPM 51
Db 21 LILLFVLLILSLQVSSVVDQSLKSVQDKVLLPRVNSPHDESEDIYQKHQKVLSV 80
Qy 52 CWGKGFPCMSQCTNELIRTERNVYQSSRYQLKGLNKGVSLLIKNVTLDHGYTC 111
Db 81 IAGK-LKVMPEYKN---RTLYDNTTY-----SLIILGLVLSRGTYS 119
Qy 112 RIQPP-GLMNDKLEL-KDIDYAKYTPAQTANGDSTTASPR-----IS---TWDEIK 184
Db 120 VQKKRGTYEVKRLALVKLSIKADFPSTPNITSGNSADTKRITCFASGFPKPRSWL 179
Qy 152 -----TLTERNGSEFQTLVTLNNNGTK-----IS---TWDEIK 184
Db 180 ENGRLEPGINTTISQDESELEYTSSQDPNTTNNHTIKCLIKGDAHVSDFTWKPRP 239
Qy 185 DSGETIRTAIHIGVVSAGLTLAIIGVLIKWSCKKK 223
Db 240 DPPSKNTLVLFAGFGAVITVVV--VVIKCF-CGR 275

RESULT 9
CD33_MOUSE STANDARD; PRT; 403 AA.
AC 063964; O63997;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Myeloid cell surface antigen CD33 precursor (Siglec-3).
GN CD33.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 33-A AND 33-B).
RC STRAIN=BALB/c; TISSUE=Bone marrow;
RX MEDLINE=94250900; PubMed=8193354;
RA Tchillian B.Z., Beverley P.C., Young B.D., Watt S.M.;
RT "Molecular cloning of two isoforms of the murine homologue of the
  myeloid CD33 antigen.",
  Blood 83:3188-3198 (1994).

```



CC -1- FUNCTION: Putative adhesion molecule of myelomonocytic-derived  
CC cells that mediates stialic-acid dependent binding to cells.  
CC Preferentially binds to alpha2,6-linked stialic acid (by  
CC similarity). The stialic acid recognition site may be masked by cis  
CC interactions with stialic acids on the same cell surface.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=33-B;  
CC IsoId=O63994-1; Sequence=Displayed;  
CC Name=33-A;  
CC IsoId=O63994-2; Sequence=VSP\_002534;  
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. SINGLE  
CC (STATIC ACID BINDING IG-LIKE LECTIN) SUBFAMILY.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: S71345; AAB30842.1; -  
CC EMBL: S71403; AAB30843.2; -  
CC MGJ: MGJ:99440; CA33.  
CC InterPro: IPR007110; IG-like.  
CC InterPro: IPR003599; IG.  
CC InterPro: IPR003006; IG\_MHC.  
CC Pfam: PF00047; IG\_2.  
CC SMART: SM00409; IG\_1.  
CC DR: Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;  
DR: Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;  
KW: Immunoglobulin domain; Alternative splicing.  
FT SIGNAL 1 16  
FT CHAIN 17 403 MYELOID CELL SURFACE ANTIGEN CD33.  
FT DOMAIN 18 240 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 241 267 POTENTIAL.  
FT DOMAIN 268 403 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 17 120 IG-LIKE V-TYPE.  
FT DOMAIN 145 228 IG-LIKE C2-TYPE.  
FT DISULFID 36 169 BY SIMILARITY.  
FT DISULFID 41 100 BY SIMILARITY.  
FT DISULFID 163 212 BY SIMILARITY.  
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 287 403 ROEATSYNHCUSPTASDAVTEGCSIHRLAISRCTALLR  
IOPIYRTHLRNKAVALTLPFWISWEGSLASTORSKCTLC  
SPVKNLCPIPLMIPVNSCIPILPEMWMLCVSLTLS -> AH  
OQDSKVSHPNENRPLQKDSPOECSVHTKISIDFWGGRPO  
EYSEI (in isoform 33-A).  
FT FTID=VSP\_002534.  
SQ SEQUENCE 403 AA; 44824 MW; F1FE6D5C393F0F1 CRC64;  
Query Match 8.2%; Score 120.5; DB 1; Length 403;  
Best Local Similarity 23.7%; Pred. No. 0.0091;  
Matches 65; Conservative 32; Mismatches 116; Indels 61; Gaps 8;

QY 155 -----TENSGSTQTLVTLHNNNGKISTMADEIKDSEGTIRTAHIGVSGA----- 203  
DB 184 LITSSTRTDSSVLTFTPOPDHGKTL-TCLVTFSGAGVATERTQLNVTKSGGMRELIV 242  
QY 204 -----LTLLALIGVILILKWYSCCKKKLS 226  
DB 243 LVAVGEATVKLLILGLCLVFLIWMCRKRTKLS 276  
RESULT 10  
ICOL MOUSE  
ID ICOL MOUSE STANDARD; PRT; 322 AA.  
AC 09JH8.  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE ICOS ligand precursor (B7 homolog 2) (B7-H2) (B7-like protein G150)  
DE (B7-related protein-1) (B7RP-1) (ICOS).  
GN ICOSL OR B7H2 OR B7RP1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.  
RC TISSUE=Lymphocytes;  
RX MEDLINE=20083495; PubMed=10617205;  
RA Yoshinaga S.K., Whoriskey J.S., Khare S.D., Sarmiento U., Guo J.,  
RA Horan T., Shih G., Zhang M., Coccia M.A., Kohno T., Tafuri-Bladt A.,  
RA Brankov D., Campbell P., Chang D., Chiu L., Dai T., Duncan G.,  
RA Elliott G.S., Hui A., McCabe S.M., Scully S., Shabinian A.,  
RA Shalee C.L., Van G., Mak T.W., Senaldi G.,  
RT "T-cell co-stimulation through B7RP-1 and ICOS."  
RL Nature 402:827-832(1999).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Thymus;  
RX MEDLINE=20015817; PubMed=10549624;  
RA Swallow M.W., Wallin J.U., Sha W.C.,  
RT "B7h, a novel costimulatory homolog of B7.1 and B7.2, is induced by  
RT TNFalpha."  
RN Immunity 11:423-432(1999).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN=C3H/HeJ; TISSUE=Fetal thymus;  
RX MEDLINE=20126021; PubMed=10657606;  
RA Ling V., Wu P.W., Finnerty H.F., Bean K.M., Spaulding V., Fouser L.A.,  
RA Leonard J.P., Hunter S.E., Zollner R., Thomas J.L., Miyashiro J.S.,  
RT "Identification of G150, a novel B7-like protein that functionally  
RT binds to ICOS receptor."  
RL J. Immunol. 164:1653-1657(2000).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC TISSUE=Peritoneal blood lymphocytes;  
RX MEDLINE=21286479; PubMed=11390480;  
RA Ling V., Wu P.W., Miyashiro J.S., Marusic S., Finnerty H.F.,  
RA Collins M.,  
RT "Differential expression of inducible costimulator-ligand splice  
RT variants: lymphoid regulation of mouse g150-b and human g150  
RT molecules."  
RL J. Immunol. 166:7300-7308(2001).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RA Ling V., Dunussi-Joannopoulos K.,  
RT "G150 molecules and uses therefor."  
RL Patent number WO0121796, 29-MAR-2001.  
CC -1- FUNCTION: LIGAND FOR THE T-CELL-SPECIFIC CELL SURFACE RECEPTOR  
CC ICOS. ACTS AS A COSTIMULATORY SIGNAL FOR T-CELL PROLIFERATION AND  
CC CYTOKINE SECRETION; INDUCES ALSO B-CELL PROLIFERATION AND  
CC DIFFERENTIATION INTO PLASMA CELLS. COULD PLAY AN IMPORTANT ROLE IN  
CC MEDIATING LOCAL TISSUE RESPONSES TO INFLAMMATORY CONDITIONS, AS  
CC WELL AS IN MODULATING THE SECONDARY IMMUNE RESPONSE BY CO-



CC STIMULATING MEMORY T CELL FUNCTION. DURING PREGNANCY, MAY FUNCTION  
CC TO SKEW THE CYTOKINE OF MATERNAL T-CELLS TOWARD IMMUNOPROTECTIVE  
CC TH2 PHENOTYPE.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=2;  
CC Comment=Additional isoforms seem to exist;  
CC Name=1;  
CC IsoId=Q9JHJ8-1; Sequence=Displayed;  
CC Name=2; Synonyms=B;  
CC IsoId=Q9JHJ8-2; Sequence=VSP\_002521;  
CC -1- TISSUE SPECIFICITY: ISOFORM 1 HIGHEST EXPRESSION IN LYMPHOID  
CC TISSUES, SUCH AS SPLEEN (MOSTLY IN THE MARGINAL ZONE), LYMPH NODES  
CC (PARTICULARLY IN THE CORTX AND IN BOTH PRIMARY AND SECONDARY  
CC FOLLICLES), THYMUS (PREDOMINANTLY IN THE MEDULLA) AND Peyer's  
CC PATCHES (MOSTLY IN THE FOLLICLES). LOWER LEVELS IN MANY  
CC NONLYMPHOID TISSUES, SUCH AS BRAIN, HEART, KIDNEY, LIVER, LUNG,  
CC SKELETAL MUSCLE AND TESTIS. PRESENT ON FRESHLY ISOLATED SPLENIC B-  
CC CELLS, T-CELLS, DENDRITIC CELLS AND MACROPHAGES. THE EXPRESSION OF  
CC ISOFORM 2 IS RESTRICTED TO HEART, SPLEEN AND KIDNEY.  
CC -1- DEVELOPMENTAL STAGE: DETECTED EARLY IN EMBRYOS; IN THE YOLK  
CC SAC AT 11.5 AND 12.5 DPC AND, TO A LESSER EXTENT, IN THE LIVER AT  
CC 14.5 DPC.  
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. BTN/MOG  
CC SUBFAMILY.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-c-type domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF216747; AAF45149.1; -  
CC EMBL; AF199027; AAF94738.1; -  
CC EMBL; AX100591; CAC36463.1; -  
CC EMBL; AX100593; CAC36464.1; -  
CC EMBL; AF394451; AAK77544.1; -  
CC MGI; MGI:1354701; Icosl.  
CC DR GO; GO:000576; C:extracellular; IDA.  
CC DR GO; GO:0016021; C:integral to membrane; NAS.  
CC DR GO; GO:0003793; F:defense/immunity protein activity; NAS.  
CC DR GO; GO:0005102; F:receptor binding activity; TAS.  
CC DR GO; GO:0045190; P:isotype switching; NAS.  
CC DR GO; GO:0042104; P:positive regulation of activated T-cell pro. . .; TAS.  
CC DR GO; GO:0045404; P:positive regulation of interleukin-4 biosyn. . .; IMP.  
CC DR GO; GO:0042110; P:T-cell activation; NAS.  
CC DR InterPro; IPR007110; Ig-like.  
CC DR InterPro; IPR003599; Ig.  
CC DR InterPro; IPR003006; Ig\_MHC.  
CC DR Pfam; PF00047; Ig\_1.  
CC DR SMART; SM00409; Ig\_1.  
CC DR PROSITE; PS50835; IG\_LIKE; 2.  
CC DR B-cell activation; Immune response; Glycoprotein;  
CC Immunoglobulin domain; Signal; Transmembrane; Multigene family;  
CC KM Alternative splicing.  
CC FT SIGNAL 1 46 BY SIMILARITY.  
CC FT CHAIN 47 322 ICOS LIGAND.  
CC FT DOMAIN 47 277 EXTRACELLULAR (POTENTIAL).  
CC FT TRANSMEM 278 298 POTENTIAL.  
CC FT DOMAIN 299 322 CYTOPLASMIC (POTENTIAL).  
CC FT DOMAIN 49 149 IG-LIKE V-TYPE.  
CC FT DOMAIN 167 263 IG-LIKE C2-TYPE.  
CC FT DOMAIN 31 38 POLY-LBU.  
CC FT DOMAIN 289 292 POLY-ALA.  
CC FT DISULFID 62 138 POTENTIAL.  
CC FT DISULFID 185 243 POTENTIAL.  
CC FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT

FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 213 213 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 321 322 HA -> TWAPVPDYLLPRYMSPLKTRGLP (in  
FT isoform 2).  
FT /FTID=VSP\_002521.  
FT R -> H (IN REF. 4 AND 5; CAC36464).  
SQ SEQUENCE 322 AA; 35960 MW; 55CCBA4AD12E47E6 CRC64;  
SQ  
Query Match 8.1%; Score 119.5; DB 1; Length 322;  
Beet Local Similarity 28.8%; Pred. No. 0.0084;  
Matches 59; Conservative 20; Mismatches 75; Indels 51; Gaps 11;  
QY 2 FSGLTNCVLLLLDOLLARSLEDGKVEGKNAVLP-----SYLPTSGTLVPMCKG 55  
DB 28 FSGIGL--FLILLSSLCASAEVEGAMVGSNVVLSICIDPHRRHNL--SGLYV----- 77  
QY 56 GFCWMSQCTNELLRTERNVY-----QKSRVQLKGDNL-----KGDVSLIKNV 101  
DB 78 ---TWQ-----INPEVSVITYLPYKSPGINVDSYKNGHLSLDSMKQGFSLYLNKV 128  
QY 102 TLDDHGYCCRIOPPGIAND-----KKLE--LKLDIKAKVTPAQTAHGDSTTASPRIT 154  
DB 129 TPQDTQEPTRCV-----FNTATLVLKILEEVRLRVANFSTPVISTSSNPQGERITY 184  
QY 155 -TERNGSETOTLVTLHNNNGTKIST 178  
DB 185 CMSKNGYPEPMLVWINTDNLIDT 209  
RESULT 11  
SIL5\_HUMAN  
ID SIL5\_HUMAN STANDARD; . PRT; 551 AA.  
AC O15389;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Sialic acid binding Ig-like lectin 5 precursor (Siglec-5) (obesity-  
DE binding protein 2) (OB binding protein-2) (OB-BP2) (CD33 antigen-like  
DE 2) (CD170 antigen).  
OS SIGLECS OR OBSP2 OR CD33L2.  
GN Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euteleota; Primates; Carnivora; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=98402360; PubMed=9731071;  
RA Cornish A.L., Freeman S., Forbes G., Ni J., Zhang M., Cepeda M.,  
RA Gentz R., Augustus M., Carter K.C., Crocker P.R.;  
RT "Characterization of siglec-5, a novel glycoprotein expressed on  
RL blood 92.2123-2132(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Erythrocytopenia;  
RX MEDLINE=99357812; PubMed=10428856;  
RA Patel N., Brinkman-Van der Linden E.C.M., Altmann S.W., Gish K.C.,  
RA Balasubramanian S., Thams J.C., Peterson D., Bell M.P., Bazan J.F.,  
RA Varki A., Kastelein R.A.;  
RT "OB-BP1/Siglec-5, A leptin- and sialic acid-binding protein of the  
RL immunoglobulin superfamily.";  
RL J. Biol. Chem. 274:22729-22738(1999).  
RN [3]  
RP ERRATUM.  
RA Patel N., Brinkman-Van der Linden E.C.M., Altmann S.W., Gish K.C.,  
RA Balasubramanian S., Thams J.C., Peterson D., Bell M.P., Bazan J.F.,  
RA Varki A., Kastelein R.A.;  
RL J. Biol. Chem. 274:28058-28058(1999).  
RN [4]

```

FT CARBOHYD 375 375 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 72 72 V -> A (in dbSNP:1973019).
FT VARIANT 215 215 /FTID=VAR_014249.
FT VARIANT 322 322 M -> V (in dbSNP:1807124).
FT VARIANT 322 322 F -> S (in dbSNP:2278831).
FT CONFLICT 309 309 E -> K (IN REF. 1).
FT CONFLICT 358 358 R -> W (IN REF. 1).
FT CONFLICT 388 388 A -> P (IN REF. 1).
FT CONFLICT 403 403 S -> N (IN REF. 1).
SQ SEQUENCE 551 AA; 60715 MW; 2FEAA2B6B341EF6AF CRC64;

Query March 8.1%; Score 118.5; DB 1; Length 551;
Best Local Similarity 23.9%; Pred. No. 0.02; Mismatches 86; Indels 55; Gaps 10;
Matches 57; Conservative 40;

QY 11 LLLLLQLLARSLEDG--YKVEGKN-----AYLPCSYLP-----TSGLVPMCMGK 55
DQ 4 LLLPLFLMGSLGQKPYEHLQVQKSVTVQGLCVLVPDCSPSYPRMSWSSPLVYVFRD 63
QY 56 GPCWMSCTMELRTD--ERNVYQKSRQYQKGLDKNGDVSLIKVATLDPHQTCRI 113
DQ 64 GEIPYVA--EVATNPDRRVRPEYQGRFRLLGDQKKNCSLSIGARWEDTSYFFRV 120
QY 114 Q-----FPGIMNDKLEL-----KLDIK-----AAKVTPTAQTANG 143
DQ 121 ERGSDVYKSVQCKNLNLEVALLIEKPIHFLPLESGRPTRLRSLGSGCAGPLTFWSW 180
QY 144 DSTTASGRPLTTERNGSETOTLVTLHNNNGTKISTWDEIKDSGTTIRPAIHGVGS 201
DQ 181 TGNALSP--LDPEPTRSSSELTLPTRPDHGTNLTC--QMKRQGAQVTTERTVQANVS 233

RESULT 12
CD4_RABIT STANDARD; PRT; 459 AA.
ID_CD4_RABIT
AC P46630;
DT 01-NOV-1995 (Ref. 32, Created)
DT 01-NOV-1995 (Ref. 32, Last sequence update)
DT 15-SEP-2003 (Ref. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Leu-3).
GN CD4.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92390370; PubMed=1518821;
RA Hague B.F., Sawasdikosol S., Brown T.J., Lee K., Recker D.P.,
RA Kindt T.J.,
RT "CD4 and its role in infection of rabbit cell lines by human
RT immunodeficiency virus type 1."
RL Proc. Natl. Acad. Sci. U.S.A. 89:7963-7967(1992).
CC -1- FUNCTION: ACCESSORY PROTEIN FOR MHC CLASS-II ANTIGEN/T-CELL
CC RECEPTOR INTERACTION. MAY REGULATE T-CELL ACTIVATION.
CC -1- SUBUNIT: ASSOCIATES WITH P56-LCK (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
```

DR EMBL; M92840; AAA3198.1; -  
 DR PIR; A46254; A46254.  
 DR HSSP; P01730; 1MER.  
 DR GO; GO:0042101; C-T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding activity; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosym. . .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007159; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro; IPR000973; CD4-TDAG.  
 DR InterPro; IPR007110; IG-1-like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG\_2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG LIKE; 1.  
 KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;  
 FT SIGNAL 1 25  
 FT CHAIN 26 459  
 FT DOMAIN 26 396  
 FT TRANSSEM 397 419  
 FT DOMAIN 420 459  
 FT DOMAIN 26 129  
 FT DOMAIN 130 208  
 FT DOMAIN 209 318  
 FT DOMAIN 319 374  
 FT CARBOHYD 299 299  
 FT DISULFID 41 113  
 FT DISULFID 329 370  
 FT LIPID 420 420  
 FT LIPID 423 423  
 SQ SEQUENCE 459 AA; 50886 MW; B323311CBDA0013D CRC64;  
 Query Match 7.9%; Score 116; DB 1; Length 459;  
 Best Local Similarity 26.8%; Pred. No. 0.026; Matches 45; Conservative 21; Mismatches 54; Indels 48; Gaps 7;  
 Db 9 CULLLLOLLARSLDEGYKVEVGNAYLPCSYTLPTSGLVPMCGKGFCEWQCTNEL 68  
 Db 9 CLLLVLPLALPLPATWGTGTVVRK-----AGAIYELP-----CQSQGRNSVF 51  
 Qy 69 RTDERNTY-----QKSSRYQLKGD-----LAKGVSLIKNTLDDHGT 109  
 Db 52 NWKHANOVKILGNQSSSSSFWLKGNSPLSNRVESKKNNMDQGFPLVIXDLRMDDSGT 111  
 Qy 110 CCRLOPGLAMNDKLEKLDIKAKVTP-AQTAGDSTTASPRITTE 156  
 Db 112 ICEV-----GDKMEVELLVRLTANDNTLLHGQSL-----TLTLE 148

RT glycoprotein gene.";  
 RL Genomics 23:36-41(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gardiner M.V., Mathieu J.M.;  
 RT "Murine and human MOG are highly conserved: cDNA analysis.";  
 RL Trans. Am. Soc. Neurochem. 24:234-234(1993).  
 RN [3]  
 RP SEQUENCE OF 29-246 FROM N.A.  
 RC STRAIN-BALB/c; TISSUE=Brain;  
 RX MEDLINE=93376728; PubMed=8367453;  
 RA Pham-Dinh D., Mattei M.-G., Nussbaum J.-L., Rousset G.,  
 RA Pontarotti P., Roedel N., Mather I.H., Artzt K., Lindahl K.F.,  
 RA Dautigny A.;  
 RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the  
 RT immunoglobulin superfamily encoded within the major  
 RT histocompatibility complex.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994(1993).  
 RN [4]  
 RP SEQUENCE OF 29-54.  
 RC STRAIN-BALB/c; TISSUE=Brain;  
 RX MEDLINE=92218912; PubMed=1373175;  
 RA Amiguet P., Gardiner M.V., Zanetta J.-P., Mathieu J.-M.;  
 RT "Purification and partial structural and functional characterization  
 RT of mouse myelin/oligodendrocyte glycoprotein.";  
 RL J. Neurochem. 58:1676-1682(1992).  
 CC -1- FUNCTION: MINOR COMPONENT OF THE MYELIN SHEATH MAY BE INVOLVED IN  
 CC COMPLETION AND/OR MAINTENANCE OF THE MYELIN SHEATH AND IN CELL-  
 CC CELL COMMUNICATION.  
 CC -1- SUBUNIT: MAY FORM HOMODIMERS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE CNS. WHERE IT IS  
 CC LOCALIZED ON THE SURFACE OF MYELIN AND OLIGODENDROCYTE CYTOPLASMIC  
 CC MEMBRANES.  
 CC -1- DISEASE: REDUCED CONCENTRATIONS OF MOG ARE OBSERVED IN JIMPY AND  
 CC QUACKING DYSPHYELINATING MUTANT MICE.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. BTN/MOG  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -1- CAUTION: DO NOT CONFUSE MYELIN-OLIGODENDROCYTE GLYCOPROTEIN (MOG)  
 CC WITH OLIGODENDROCYTE-MYELIN GLYCOPROTEIN (OMG).  
 CC -----  
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 DR EMBL; L29503; AAC42023.1; -  
 DR EMBL; L29498; AAC42023.1; JOINED.  
 DR EMBL; L29501; AAC42023.1; JOINED.  
 DR EMBL; L29501; AAC42023.1; JOINED.  
 DR EMBL; L29499; AAC42023.1; JOINED.  
 DR EMBL; L29502; AAC42023.1; JOINED.  
 DR EMBL; U64572; AAB08096.1; -  
 DR EMBL; L20942; AAA03180.1; -  
 DR PIR; A55717; A55717.  
 DR MGI; MGI:97435; Mog.  
 DR InterPro; IPR007110; IG-1-like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR InterPro; IPR003596; IG\_V.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG LIKE; 1.  
 KW Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.  
 FT SIGNAL 1 28  
 FT CHAIN 29 246  
 FT DOMAIN 29 156  
 FT TRANSSEM 157 177  
 FT DOMAIN 178 209  
 FT TRANSSEM 210 230  
 MYELIN-OLIGODENDROCYTE GLYCOPROTEIN.  
 EXTRACELLULAR (POTENTIAL).  
 CYTOPLASMIC (POTENTIAL).  
 POTENTIAL.

FT DOMAIN 231 246 - EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 31 144 IG-LIKE V-TYPE.  
 FT DISULFID 52 126 POTENTIAL.  
 FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 21 21 L -> IL (IN REF. 2).  
 FT CONFLICT 32 32 R -> G (IN REF. 4).  
 FT CONFLICT 95 95 G -> E (IN REF. 3).  
 FT CONFLICT 169 169 P -> S (IN REF. 2).  
 SQ SEQUENCE 246 AA; 28271 MM; 1F1A8A4A0D5CFE89; CRC64;  
 Query Match 7.8%; Score 114; DB 1; Length 246;  
 Best Local Similarity 27.4%; Pred. No. 0.018;  
 Matches 37; Conservative 30; Mismatches 52; Indels 16; Gaps 5;  
 QY 7 LNCVLLDQLLLA-----RSLEDGYVE--VGNAYLPGSYTLPTSGTLVPMCKGKFCP 59  
 DB 14 LSLILLLLDLSGSCYACQFRITGYPYIRALVGEALPCISPGKNATGMEVGMYS--P 71  
 QY 60 WSQCTNELLRTDERNNTYKSSRYQ-----LKGDLKGDVSLIKNVTLDHGTTCRQ 114  
 DB 72 FSRVVA--LYRNGKDDADQAPRYRGRTELLKETTISEGVTLRINQVRPSDSGYTCFPR 129  
 QY 115 PFLMNDKLELELD 129  
 DB 130 DHSYGEAMLELVE 144  
 RESULT 14  
 ID SIL9 HUMAN STANDARD; PRT; 463 AA.  
 AC Q9Y336; Q9BY19;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Static acid-binding Ig-like lectin 9 precursor (Siglec-9) (FOAP-9 protein).  
 GN SIGLEC9.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RC TISSUE=BONE marrow;  
 RX MEDLINE=2036134; PubMed=10903842;  
 RA Fousias G., Yousef G.M., Diamandis E.P.;  
 RT "Identification and molecular characterization of a novel member of the siglec family (SIGLEC9).";  
 RL Genomics 67:171-178(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20357333; PubMed=10801862;  
 RA Zhang J.Q., Nicoll G., Jones C., Crocker P.R.;  
 RT "Siglec-9, a novel static acid binding member of the immunoglobulin superfamily expressed broadly in human blood leukocytes.";  
 RL J. Biol. Chem. 275:22121-22126(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-120.  
 RC TISSUE=Periplheral blood;  
 RX MEDLINE=20357334; PubMed=10801860;  
 RA Angata T., Varki A.;  
 RT "Cloning, characterization, and phylogenetic analysis of siglec-9, a new member of the CD33-related group of siglecs. Evidence for co-evolution with static acid synthesis pathways.";  
 RL J. Biol. Chem. 275:22127-22135(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND VARIANT GLU-315.  
 RA Takeyama K., Fujii Y., Turfani K., Naitou K., Kawaguchi A., Uka Y., Amemura C., Yajima Y., Yazaki M.;  
 RT "Molecular cloning of a novel gene, FOAP-9, which are induced by oxylized LDL in human macrophages";  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [5]

RP SEQUENCE FROM N.A.  
 RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V., Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S., Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Gares J., Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S., Andreise T., Trankheim M., Attix C., Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B., Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;  
 RT "Sequence analysis of chromosome 19q13.4.";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Putative adhesion molecule that mediates static-acid dependent binding to cells. Preferentially binds to alpha2,3- or 2,6-linked static acid. The static acid recognition site may be masked by cis interactions with static acids on the same cell surface.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Expressed by peripheral blood leukocytes (neutrophils and monocytes but not eosinophils). Found in liver, fetal liver, bone marrow, placenta, spleen and in lower levels in skeletal muscle, fetal brain, stomach, lung, thymus, prostate, brain, mammary, adrenal gland, colon, trachea, cerebellum, testis, small intestine and spinal cord.  
 CC -1- DOMAIN: Contains an intracytoplasmic motif referred as immunoreceptor tyrosine-based inhibitor motif (ITIM). This motif is involved in downmodulation of cellular functions as the termination of the immune response.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. SIGLEC (STATIC ACID BINDING IG-LIKE LECTIN) SUBFAMILY.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF135027; AAD26428.2; -;  
 DR EMBL; AF247180; AAF87223.1; -;  
 DR EMBL; AF227924; AAF71455.1; -;  
 DR EMBL; AB026265; BAB41100.1; -;  
 DR EMBL; AC011473; AAG23261.1; -;  
 DR Genew; HGNC:10878; SIGLEC9.  
 DR MIM; 605640; -;  
 DR GO; GO:0005887; C:integral to plasma membrane; NAS.  
 DR GO; GO:0005530; F:lectin; NAS.  
 DR GO; GO:0007166; F:cell surface receptor linked signal transdu. .; NAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig\_3.  
 DR SMART; SM00409; IG; 2.  
 DR PROSITE; PSS0835; IG\_LIKE; 3.  
 KW Cell adhesion; Lectin; Antigen; Transmembrane; Signal; Glycoprotein; Immunoglobulin domain; Repeat; Polymorphism.  
 KW SIGNAL 1 17  
 FT CHAIN 18 463  
 FT DOMAIN 18 348  
 FT TRANSMEM 349 369  
 FT DOMAIN 370 463  
 FT DOMAIN 20 140  
 FT DOMAIN 146 229  
 FT DOMAIN 236 336  
 FT SITE 431 436  
 FT SITE 454 459  
 FT DISULFID 36 170  
 FT DISULFID 41 102  
 FT DISULFID 164 213  
 FT DISULFID 272 320  
 FT CARBOHYD 101 101  
 N-LINKED (GLCNAC. .) (POTENTIAL).  
 POTENTIAL.  
 EXTRACELLULAR (POTENTIAL).  
 POTENTIAL.  
 CYTOPLASMIC (POTENTIAL).  
 IG-LIKE V-TYPE.  
 IG-LIKE C2-TYPE 1.  
 IG-LIKE C2-TYPE 2.  
 ITIM MOTIF.  
 SLAM-LIKE MOTIF.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.

```

FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 100 100 K-> E (in dbSNP:2075803).
FT VARIANT 125 125 S-> N (in dbSNP:200658).
FT VARIANT 147 147 /FTid=VAR_014255.
FT VARIANT 147 147 N-> K (in dbSNP:273687).
FT VARIANT 315 315 /FTid=VAR_014256.
FT VARIANT 316 316 A-> E (in dbSNP:2258983).
FT VARIANT 316 316 /FTid=VAR_014257.
FT MUTAGEN 120 120 A-> D (in dbSNP:273688).
FT CONFLICT 269 269 R-> K; LOSS OF SIMILIC ACID BINDING.
SQ SEQUENCE 463 AA; 50081 MM; 807BDCE0D18191F2 CRC64;

```

```

Query Match 7.7%; Score 112.5; DB 1; Length 463;
Best Local Similarity 21.6%; Pred. No. 0.051;
Matches 55; Conservative 42; Mismatches 95; Indels 63; Gaps 10;

QY 10 VLLHLLQLLAR-----SLDDGYKVEGKNAVLPSCSYTLPTSGTLVPMCMGKGF 58
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 LLLLLPLLMKRRERAGCOTSKLTMQSSVTVEGLCVHPCFSFSPSHWITPGVNAVY- 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 59 PMSOCTNELLRTDE-----RNVYQKSSRYOLKGDINKDVSLIKNVTLDHGT 108
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 -WFR---EGANTDDAPVATNPPARAWEIEFRDRFHLGLDHTKXCTLSIDARSDAGR 116
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 109 YCCIIQFPGI-MNKKLEKIDIKAKTPTAOTAHGSDTSSPRTLT----- 154
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 117 YFFEMEKSGIKMVKHRLSNVTALTTPRPILPGLTESCCPONTLCISVWACEQGTTP 176
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 155 -----TERNGSETQTLVTLHNNGTKISTWDEKDSGETIRT--AIHIG 197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 177 MISWIGTSVSLDSDSTR--SSVTLLIPQPDHGTSLTC--QVTFPGASVTTNKTYHLN 231
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 198 VGV-SAGITLALIG 211
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 232 VSVPPQNLVMTVFG 246
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 15
CD4_SAISC STANDARD; PRT; 457 AA.
ID_CD4_SAISC
AC 029037;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
   T4/Leu-3).
GN CD4.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RP SEQUENCE FROM N.A.
RA Tatsumi M., Hashimoto O.;
RU Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACCESSORY PROTEIN FOR MHC CLASS-II ANTIGEN/T-CELL
CC RECEPTOR INTERACTION. MAY REGULATE T-CELL ACTIVATION.
CC -1- SUBUNIT: ASSOCIATES WITH P56-LCK (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

```

CC -----
DR EMBL; D86588; BAA1311.1; -.
DR HSSB; P01730; 1MBR.
DR GO; GO:0042101; C: T-cell receptor complex; ISS.
DR GO; GO:00015026; F: coreceptor activity; ISS.
DR GO; GO:0042289; F: MHC class II protein binding activity; ISS.
DR GO; GO:0006955; P: immune response; ISS.
DR GO; GO:0045086; P: positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P: T-cell differentiation; ISS.
DR GO; GO:0045058; P: T-cell selection; ISS.
DR GO; GO:0007169; P: transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro; IPR000973; CD4_TcRg.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25 BY SIMILARITY.
FT CHAIN 26 457 T-CELL SURFACE GLYCOPROTEIN CD4.
FT TRANSSEM 396 417 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 418 457 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 126 202 IG-LIKE V-TYPE.
FT DOMAIN 203 316 IG-LIKE C2-TYPE 1.
FT DOMAIN 317 373 IG-LIKE C2-TYPE 2.
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 41 109 BY SIMILARITY.
FT DISULFID 155 184 BY SIMILARITY.
FT DISULFID 327 369 BY SIMILARITY.
FT LIPID 418 418 PALMITATE (BY SIMILARITY).
FT LIPID 421 421 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 457 AA; 50871 MM; 57EBD6344005A015 CRC64;

```

```

Query Match 7.5%; Score 109.5; DB 1; Length 457;
Best Local Similarity 24.9%; Pred. No. 0.09;
Matches 46; Conservative 26; Mismatches 80; Indels 33; Gaps 8;

QY 1 MFSGTLNLCVLLQLLARSLEDGYKVEGKN---AYLPSCSYTLPTSGTLVPMCMGKGF 57
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MNGGIPRHLVLTQLLPLPVHTGKTAVLQKGEVVELPETSLSKN--VPRHWKTS- 56
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 58 CPWSQ-----CTNELRTDERNVYQKSSRYOLKGDINKDVSLIKNVTLDHGTCCR 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 57 ---DQIKILGQNVFVRGSGSKLDRIDSK---KSSMDRGSFPLLDARIEDSETVCE 110
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 113 IQFGLMNDKKLEKIDIKAKTPTAOTAHGSDT-----ASPRTLTERNGSETQ 163
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 111 VE-----SKKEVELQVFGITANPDTHLLOGSLTLLTESPPSSSVECTSPRGKRIR 164
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 164 TLVTL 168
DB 165 GRKTL 169

```

Search completed: November 22, 2003, 05:41:03  
 Job time : 16.5189 secs

THIS PAGE BLANK (USPTO)



QY 61 SOCTNELRTDERNTYOKSSRYQLKGDINKGVSLLIKNVTLDHGTCCRIQPFGLMN 120  
 DB 61 SOCTNELRTDERNTYOKSSRYQLKGDINKGVSLLIKNVTLDHGTCCRIQPFGLMN 120  
 QY 121 DKLEIKLIDIKAKVTPAQTAGDSTTASPRITLTERNGSETOVLTLNNNGTKISTWA 180  
 DB 121 DKLEIKLIDIKAKVTPAQTAGDSTTASPRITLTERNGSETOVLTLNNNGTKISTWA 180  
 QY 181 DEIKDSGETIRTAIHIGVVSAGLTALIIIGVILIKWYCKKKKSLSLITLANLPFG 240  
 DB 181 DEIKDSGETIRTAIHIGVVSAGLTALIIIGVILIKWYCKKKKSLSLITLANLPFG 240  
 QY 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281  
 DB 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281

## RESULT 2

Q96K94 PRELIMINARY; PRT; 301 AA.  
 ID Q96K94;  
 AC Q96K94;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Hypothetical protein FLJ14428 (Putative kidney injury molecule-3).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Nabeura T., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y., Oshima A.;  
 RT "NEO human cDNA sequencing project";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Hepatosoma;  
 RA Kuenen E.W., Ichimura T., Bonventre J.V.;  
 RT "A homolog to human kidney injury molecule-1 is expressed in hepatoma cells";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK027334; BAB55044.1; -;  
 DR EMBL; AY069944; AAL55401.1; -;  
 DR InterPro; IPR007110; IG-like.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 301 AA; 33351 MW; 4877B7D9E40A4161 CRC64;

Query Match 60.1%; Score 883; DB 4; Length 301;

Best Local Similarity 63.7%; Pred. No. 1.1e-67;  
 Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSGTLNLCVLLLLLOLLARSLSDGYKVEGNAYLPCSYTLPTSGTLVPMCMGKGFPCW 60  
 DB 1 MFSGTLNLCVLLLLLOLLARSLSDGYKVEGNAYLPCSYTLPTSGTLVPMCMGKGFPCW 60  
 QY 61 SOCTNELRTDERNTYOKSSRYQLKGDINKGVSLLIKNVTLDHGTCCRIQPFGLMN 120  
 DB 61 SOCTNELRTDERNTYOKSSRYQLKGDINKGVSLLIKNVTLDHGTCCRIQPFGLMN 120  
 QY 121 DKLEIKLIDIKAKVTPAQTAGDSTTASPRITLTERNGSETOVLTLNNNGTKISTW 179  
 DB 121 DKLEIKLIDIKAKVTPAQTAGDSTTASPRITLTERNGSETOVLTLNNNGTKISTW 179  
 QY 181 DEIKDSGETIRTAIHIGVVSAGLTALIIIGVILIKWYCKKKKSLSLITLANLPFG 240  
 DB 181 DEIKDSGETIRTAIHIGVVSAGLTALIIIGVILIKWYCKKKKSLSLITLANLPFG 240

DB 180 ANELRDSRLANDLRSGATIRIGIYGAGLCAGLALILFGLIFKWSHSEKIQNLSTL 239

QY 231 ITLANLPFGGLANAGAVRIRSEENITYTEENYEVENSNEYCYVNS-QOPS 281

DB 240 ISLANLPFGGLANAGAVRIRSEENITYTEENYEVENSNEYCYVNS-QOPS 291

## RESULT 3

Q8TD00 PRELIMINARY; PRT; 301 AA.  
 ID Q8TD00;  
 AC Q8TD00;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Hepatitis A virus cellular receptor 2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Zhang W., Wan T., Li N., Cao X.;  
 RT "Novel human hepatitis A virus cellular receptor";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF251707; AAM19100.1; -;  
 DR Genew; HGNC:18437; HAVCR2.  
 DR InterPro; IPR003599; IG-  
 DR InterPro; IPR007110; IG-like.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Receptor.  
 SQ SEQUENCE 301 AA; 33363 MW; 7DAEA13F01D3587B CRC64;

Query Match 59.7%; Score 877; DB 4; Length 301;

Best Local Similarity 63.4%; Pred. No. 3.7e-67;  
 Matches 185; Conservative 26; Mismatches 69; Indels 12; Gaps 4;

QY 1 MFSGTLNLCVLLLLLOLLARSLSDGYKVEGNAYLPCSYTLPTSGTLVPMCMGKGFPCW 60  
 DB 1 MFSGTLNLCVLLLLLOLLARSLSDGYKVEGNAYLPCSYTLPTSGTLVPMCMGKGFPCW 60  
 QY 61 SOCTNELRTDERNTYOKSSRYQLKGDINKGVSLLIKNVTLDHGTCCRIQPFGLMN 120  
 DB 61 SOCTNELRTDERNTYOKSSRYQLKGDINKGVSLLIKNVTLDHGTCCRIQPFGLMN 120  
 QY 121 DKLEIKLIDIKAKVTPAQTAGDSTTASPRITLTERNGSETOVLTLNNNGTKISTW 179  
 DB 121 DKLEIKLIDIKAKVTPAQTAGDSTTASPRITLTERNGSETOVLTLNNNGTKISTW 179  
 QY 180 ANELRDSRLANDLRSGATIRIGIYGAGLCAGLALILFGLIFKWSHSEKIQNLSTL 239  
 DB 180 ANELRDSRLANDLRSGATIRIGIYGAGLCAGLALILFGLIFKWSHSEKIQNLSTL 239  
 QY 231 ITLANLPFGGLANAGAVRIRSEENITYTEENYEVENSNEYCYVNS-QOPS 281  
 DB 240 ISLANLPFGGLANAGAVRIRSEENITYTEENYEVENSNEYCYVNS-QOPS 291

## RESULT 4

Q8WM60 PRELIMINARY; PRT; 142 AA.  
 ID Q8WM60;  
 AC Q8WM60;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Similar to hypothetical protein FLJ14428.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;





```

Db      11 LLLLLPGSV-DEYEYVKGAVVGHVPVITPCITYS--TRGIGITTCWGRGQCPFPSSCQNTILWT 67

Qy      71 DERANTYOKSSAPYOKAGDINKGDVSLIKNVLLDHDGTCYCCRCQFGLMADKLELKD1 130
      68 NGYQVYTHSSGRINIKGRISBEGVSLTIENSVDSDGLCCRCVEIIPGEMFDQMTSLEY 127

Qy      131 KAAKVT-PAQTAHGDSTTASPRLLTTEBNGSEFQTLVT-----L 168

Db      128 KPEIETSPETPTTTPTRPTTPTLTISTRTHVFTSRVSTSTPEQOTHNKEITTFYA 187

Qy      169 HNNNGCKIST-----WADIEMDGE-----TITPAI-----HIGVSGAGLT 205

Db      188 HETTALEVETPSPYTPADWNGVYTSSEEMANNHTVRIPLRKQPNPTKGFVNGSVAA-LL 246

Qy      206 LALITGVLILKMYSCKKKGLSLISLITLANLPFGGLANCAVIRSEENIYITTE 260

Db      247 LLLASTVAVTVYIIIRKKGSGISFPAFVSKSRALQONAIYHPRADENIYIIED 301

```

ID	Q8VIM2	PRELIMINARY	PRT	305 AA.
AC	Q8VIM2			
DT	01-MAR-2002	(TREMBLrel. 20, Created)		
DT	01-MAR-2002	(TREMBLrel. 20, Last sequence update)		
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)		
DE	TIM1.			
GN	TIM1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxId=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BAJ/c; TISSUE=Spleen;			
RA	McIntire J.J., Umetsu S.E., Akbari O., Potter M., Barsh G.S.,			
RA	Freeman G.J., Umetsu D.T., Dekruyff R.H.;			
RT	"Tapi, a major T cell regulatory locus that controls the development			
RT	of airway hyperreactivity, cosegregates with variants in a novel gene			
RL	family.";			
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF399829; AAL35774.1; -			
DR	MGD; MGI:2159680; Timd1.			
DR	InterPro; IPR003599; IG.			
DR	InterPro; IPR007110; IG-like.			
DR	InterPro; IPR003006; IG_MHC.			
DR	Pfam; PF00047; Ig_1.			
DR	SMART; SM00409; Ig_1.			
DR	PROSITE; PS50835; IG_LIKE; 1.			
SQ	SEQUENCE 305 AA; 33391 MW; 8FAEA38627FBE85FB CRC64;			
Query Match 24.8%; Score 364.5; DB 11; Length 305;				
Best Local Similarity 31.2%; Pred. No. 3.8e-23;				
Matches 92; Conservative. 39; Mismatches 111; Indels 53; Gaps 7				
QY	16 LILARSLENGY---KVEVGKNAVLP	CSYTLPTSGTLVPMCMGKGF	CPMSQCTNELRLTDE	72
DB	12 ILLLEGYATDSYVYGVGVH	RPVTLPCYTS--TYRGLTTTCMKRG	CCPSAACNTLWING	69
QY	73 RNVTYQKSSRGYOLKGD	LNKGVSLIIKVNTLDDHGTTCR	IQPGILNDKYLELKDIX	131
DB	70 HNVTVQKSSRYNLKGI	ISEGVDLSLTLENSVESDSGL	YCCRAVEIPGFVNDQKTVF	129
QY	132 -----AAKVP	PAQTANHDSVTAS-----BRLT	LTENSEGTQTLVTLHNNGTKI--	176
DB	130 ELPTRPPTTRPTTR	PAIGRPPTIISRSRTHVPTSIK	SVSTPPTSTHTWTHKRP	189
QY	177 -----	STWDEIKDSGETIRT-----	AIHIGVVSAGLT	205
DB	190 HETTA	EVGICPSBHTPTDNGVLTSSG	DTSMNTEALPPKPKQKPTK	248
QY	206 IALLIGVLLKMY	SCKKKLSSLSLITLANLP	PGGLANNAVAVRINSENIYITEE	260

Db	249	LLLVSVATLTRYILMKRKSASLSVAFRVSKTEALQNAVAHVSRAEDNIYYED	303
RESULT 8			
ID	043656	PRELIMINARY;	PRT; 359 AA.
AC	043656		
DT	01-JUN-1998	(TREMblrel, 06, Created)	
DT	01-JUN-1998	(TREMblrel, 06, Last sequence update)	
DT	01-MAR-2003	(TREMblrel, 23, Last annotation update)	
DT	Hepatitis A virus cellular receptor 1.		
GN	HVACR-1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver;		
RA	MEDLINE=96325180; Pubmed=9658108;		
RA	Felgelsstock D., Thompson P., Mattoo P., Zhang Y., Kaplan G.G.;		
RT	"The human homolog of HAVcr-1 codes for a hepatitis A virus cellular		
RT	receptor ";		
RL	J. Virol. 72:6621-6628(1998).		
DR	EMBL, AF043724; AAC93862.1; -.		
DR	InterPro; IPR003599; IG.		
DR	InterPro; IPR007110; IG_1like.		
DR	InterPro; IPR003006; IG_MHC.		
DR	Pfam; PF00047; IG; 1.		
DR	SMART: SMO0409; IG; 1.		
DR	PROSITE; PS50835; IG_LIKE; 1.		
DR	PROSITE; PS00290; IG_MHC; 1.		
KW	Receptor.		
SQ	SEQUENCE 359 AA; 38704 MW; C207PEC562DC62CA CRC64;		
Query Match	22.4%;	Score 328.5;	DB 4; Length 359;
Best Local Similarity	27.8%;	Pred. No. 5.9e-20;	
Matches 102;	Conservative 49;	Mismatches 93;	Indels 123; Gaps 11;
QY	10	VLLILQLILARSLDGYKV--EVGKNAVLPQSYTLPTSGTLVPMCMGKGFCEPMSCQTNEL	67
DB	6	VILILHLHLDVSVAGSVKVGEGAPSTYLCHY----SGAVTSMCMNRGSCSLFTCONGI	61
QY	68	LRTDERNVYQKSSRYQLKGDNLKGDVSLIKNVTLDDHGTYCCRIQFPGILMDKLELK	127
DB	62	VWTGTHVYTRKDRYKLGDLSPRDVSLTIENTAVSDSGVCCRVHRRGMFNDMKITVS	121
QY	128	LDIQAAYT-----	136
DB	122	LEIYPPKVTTPPIVTVPTVATSTVPTTTPPTTVPTTWSIPTTTVPTMTVST	181
QY	137	-----PAQAHGDSFTAS-----PRTL-----	153
DB	182	TTSVPTTSTPTTTSVTVTTTTSFVFPMPPLRQNHPEVATSPSPQPAETHPTLGLAI	241
QY	154	-----TTERNGSETQLVTLANNNGTKISTMADEIKDSGETIRT-AIHIGVG	199
DB	242	RREPSSPIVSYTTDGDNDVTBESDGLMNNNQTLF-----LBSLTLTANTTKGIYAGVC	296
QY	200	VSAGLTALLIIGVILIKWYSCCKKLSLSLITLANIPPGGLANAGVRIKSENIYITIE	259
DB	297	ISV-LVILALGVIIAKYFF-KKEVQOLS-VSPSSLIQIKALQNAVEKEVOAEDNIY-IE	352
QY	260	ENVEVE 266	
DB	353	NSLYATD 359	
RESULT 9			
ID	08VBWO	PRELIMINARY;	PRT; 305 AA.
AC	08VBWO;		

DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE TIMD2.  
 GN TIMD2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DB/2, and BALB/c; TISSUE=Spleen;  
 RA Mcintire J.J., Umetsu S.E., Akbari O., Potter M., Barch G.S.,  
 RA Freeman G.J., Umetsu D.T., Dekruyff R.H.;  
 RT "Tadp, a major T cell regulatory locus that controls the development  
 of airway hyperreactivity, cosegregates with variants in a novel gene  
 family.";  
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF39828; AAL35773.1; -  
 DR EMBL; AF39827; AAL35772.1; -  
 DR MGD; MGI:2159681; Timd2.  
 DR InterPro; IPR003599; IG\_1.  
 DR InterPro; IPR007110; IG\_1-like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG\_1.  
 DR SMART; SM00409; IG\_1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR SMART; PS50835; IG\_LIKE; 1.  
 SQ SEQUENCE 305 AA; 33519 MW; 68E69370FE23C9D0 CRC64;

Query Match 22.2%; Score 326; DB 11; Length 305;  
 Best Local Similarity 29.3%; Pred. No. 7.7e-20;  
 Matches 87; Conservative 49; Mismatches 117; Indels 44; Gaps 8;

QY 14 LQLLARLENGYKYE--VGKNAVIPCSTYLTPTSGTLVPMCKGKFCWSCQTNELRTD 71  
 DB 11 LILLPGAVESHATAVQGLAGHPVTLPCITYSTHLGQ-IVPMCKGAGECHSYCIRSLIWTN 69  
 QY 72 ERNVTYQKSSRYOLKGDINKGVSLLIKNVTLDHGTCCRIQFGLMND--DKKLELKD 129  
 DB 70 GYTLTHQNSRYOLKGNISSEGNVSLTIENTVGDGPRCCVETIGAHFPMYLEVPE 129  
 QY 130 IK-----AAKVTPAQTAHGDSTASPTLTTERNGS----- 161  
 DB 130 ISTSPPTPTATGRPTTISTRTSTHVTSTRTSTSTSPATETERYKPEATTFYPDTAAE 189  
 QY 162 -TQTLVTL-----HNNNGTKISTWADI-----KSGETIRRAIHGVGSAGLTALIG 211  
 DB 190 VETLTPSPADWHTNVTSSDDPMDNTTEVIPPOKQKNLNGFYGISIAA-LILMLLS 248  
 QY 212 VLILKWKCKKKKSLSLITLANLPQGLANAGAVIRSENIYTIENYEVENS 268  
 DB 249 TMTVITRYVMKRSLSLFAVAFPIKISKASPKKVERTRCDQVYIIEDTTPPEBS 305

## RESULT 10

Q96D42 PRELIMINARY; PRT; 364 AA.

AC Q96D42.  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 NC NCI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Straube R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC013325; AAH13325.1; -  
 DR Genew; HGNC:17866; HAVCR1.

DR InterPro; IPR003599; IG\_1.  
 DR InterPro; IPR007110; IG\_1-like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG\_1.  
 DR SMART; SM00409; IG\_1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 364 AA; 39249 MW; AA67C7DC7FAC81F1 CRC64;

Query Match 22.2%; Score 326; DB 4; Length 364;  
 Best Local Similarity 27.4%; Pred. No. 9.8e-20;  
 Matches 102; Conservative 49; Mismatches 93; Indels 128; Gaps 11;

QY 10 VLLIQLLARLENGYKYE--VGKNAVIPCSTYLTPTSGTLVPMCKGKFCWSCQTNEL 67  
 DB 6 VILSLIHLADSVAGSVKVGAGPSVTLPCY---SGAVTSCKWNKSGSLFTCCNGI 61  
 QY 68 LRTDERNVTYQKSSRYOLKGDINKGVSLLIKNVTLDHGTCCRIQFGLMNDKKLEL 127  
 DB 62 VMTNGTHTYAKDKRYLGLDLSRDLSTIENTAVSDSGVCCRVHRGMFNDMKIIVS 121  
 QY 128 LDIKAAYT----- 136  
 DB 122 LEIVPKVTTPIVTVPTVTVRTSTVPTTTPVPTTTPVPTTTPVPTTTPVPTT 181  
 QY 137 -----PAQTAHGDSTAS-----PRT 152  
 DB 182 MTSVSTTSVPTTTSIPTTTSVPTTTSVPTTTPVPTTTPVPTTTPVPTTTPVPTT 241  
 QY 153 L-----TTERNGSETQTLVTLHNNNGTKISTWADIKSGETIRT--AI 194  
 DB 242 LQGAIRREPTSSPLXYSTTDNDQVTESSDDLMNNQQLF-----LEHSLTLANTTGI 296  
 QY 195 HIGVGSAGLTALIIIGVLILKWKCKKKKSLSLITLANLPQGLANAGAVIRSEEN 254  
 DB 297 YAGCISY-LVLLALLGVIIIAKKYF--KKEVQQLS-VSFSLQIKALONAVEKEVQADN 353  
 QY 255 IYTIENYEVYE 266  
 DB 354 IY-IENSLIATD 364

## RESULT 11

Q8R183 PRELIMINARY; PRT; 305 AA.

AC Q8R183.  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical 33.5 kDa protein.  
 GN TIMD2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Straube R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC025096; AAH25096.1; -  
 DR MGD; MGI:2159681; Timd2.  
 DR InterPro; IPR003599; IG\_1.  
 DR InterPro; IPR007110; IG\_1-like.  
 DR SMART; SM00409; IG\_1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 305 AA; 33504 MW; 4105D75D95F541AB CRC64;

Query Match 21.7%; Score 318; DB 11; Length 305;  
 Best Local Similarity 28.3%; Pred. No. 3.8e-19;  
 Matches 84; Conservative 48; Mismatches 121; Indels 44; Gaps 7;

```

QY 14 LQLLARSLEDGKYVE--VGNAYLPCSYTLPTSGTLVPMCKGKGFPCMSQCTNELLRD 71
DB 11 LILLPGAVESHATVAGLAGHPVTLPCISTHLGG--IVPMCGMLGCCRSHYCIIRSLIMTN 69
QY 72 ENRVYQKSRVQLKGDNLKNGVSLIKNVTLDDHGTCCRIQFPGIMN--DKKLEKLD 129
DB 70 GTTVTHQRNSLVQKKNISEGVNSLTIENTVVDGPGVCCVAFIPAFHVDYMLEVKE 129
QY 130 IK-----AAKVTFA--QTAHGDSTASPRITLITE 156
DB 130 ISTSPTRPTATGRPTTISTRTHVPTSTRVSTSTPTPAHTETVKEPATTFYDQDTAE 189
QY 157 RKGSEIQTLLTNNGTKISTWADEI-----KDSEITRTAIHIGVGSAGITLALII 211
DB 190 VTETLPDTPADWNTVNTSSDDPMDNTAEVIPPQPKPKNKNKGFVGISIAA--LLIIMLS 248
QY 212 VLIKKVYSCKKKLSLSLITLANLPPGGLANAGAVRIRSEENIYTIENEVEVENS 268
DB 249 TWVITRYVMKRSSELSFVAPFIRSKIGASPKKVVERTCEQDQVYIIEDTPYEEBS 305

RESULT 12
Q96H15 PRELIMINARY; PRT; 378 AA.
AC 096H15;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008988; AA008988.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PR00047; ig_1.
DR SMART; SM00409; IG_1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Hypothetical protein.
SQ SEQUENCE 378 AA; 41550 MW; 758B9E930C82BFESD CRC64;

Query Match 18.7%; Score 274.5; DB 4; Length 378;
Best Local Similarity 42.2%; Pred. No. 2.7e-15;
Matches 57; Conservative 22; Mismatches 51; Indels 5; Gaps 2;

QY 30 VGNALPLPSYTLPTSGTLVPMCKGKGFPCMSQCTNELLRDTRNVYQKSRVQLKGD 89
DB 32 LGHRVTLPLCLYS-SWHSNMSKMGCDQCPYSGCKALIRITDMRTSRKSAKYRLOGIT 90
QY 90 NKGDVSLIKNVTLDDHGTCCRIQFPGIMNKKLEKLDIKAKVTPAQTAFHGDSTVAS 149
DB 91 PGGDVSLITLNPESDSGVYCCRIEVPWFMDVKNVRLNLQASATTHRTA-----TTT 146
QY 150 PRTLTERRNGSEIQT 164
DB 147 RRTTTSPTTTRQMT 161

RESULT 13
Q8CIC7 PRELIMINARY; PRT; 362 AA.
AC 08CIC7;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

```

```

DE Similar to hypothetical protein BC008988.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclitognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC032879; AA032879.1; -.
KW Hypothetical protein.
SQ SEQUENCE 362 AA; 39846 MW; 29846B68C9822685 CRC64;

Query Match 18.5%; Score 271.5; DB 11; Length 362;
Best Local Similarity 27.1%; Pred. No. 4.6e-15;
Matches 84; Conservative 46; Mismatches 103; Indels 77; Gaps 9;

QY 11 LLLLDL-----LLARSLDGYKVEVGNAYLPCSYTLPTSGTLVPMCKGKGFPCMSQ 62
DB 24 LILLWLVELMWLYLTPAASEDTIIGLQGPVTLPCYH--LSWSQSRNSKMGKSCPSNK 82
QY 63 CTNELLRDTRNVYQKSRVQLKGDNLKNGVSLIKNVTLDDHGTCCRIQFPGIMNDK 122
DB 83 CNAELLRDTRNVYQKSRVQLKGDNLKNGVSLIKNVTLDDHGTCCRIQFPGIMNDK 142
QY 123 KLEKLDIKAKVTPAQTAFHGDSTASPRITLITE 157
DB 143 KKNVLELRRAATTKKPTTTPPTTPVTTTPPELPTVTTSTVLPPTTPPQTLATTA 202
QY 158 NGEIQTLLTNNGTKISTWADEIKDSEITRTAIHIGVGSAGITLALII 217
DB 203 ESTAVTTCSTTPGSPQETTGSAFTTESLTPASN----- 240
QY 218 YSCKKKLSLSLIT--LANLPPGGLANAGAV--RIRSEENIYTIENEVEVENS 269
DB 241 -----SQRSMTITSDIAVLARPTG--SNPILSTSQLTQKTLTTSLSLQKTTSH 291
QY 270 EYCYVNSQ 279
DB 292 Q-----INSRQ 297

RESULT 14
Q95144 PRELIMINARY; PRT; 451 AA.
AC 095144;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE HAVET-1 protein precursor.
OS Cercopithecus aethiops (Green monkey) (Givet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA MEDLINE=97015129; PubMed=8861957;
RA Kaplan G., Totsuka A., Thompson P., Akatsuka T., Moritsugu Y.,
RA Feinstein S.M.;
RT "Identification of a surface glycoprotein on African green monkey
RT kidney cells as a receptor for hepatitis A virus.";
RL EMBL J. 15:4282-4296(1996).
DR EMBL; X98252; CAA6906.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR SMART; SM00409; IG_1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Signal.

```

FT SIGNAL 1 17 POTENTIAL.  
 SQ SEQUENCE 451 AA; 48774 MM; 5D395C5455AA4332 CRC64;  
 Query Match 18.2%; Score 267.5; DB 6; Length 451;  
 Best Local Similarity 21.7%; Pred. No. 1.4e-14;  
 Matches 97; Conservative 51; Mismatches 96; Indels 203; Gaps 11;

QY 7 LNCVLLQLLLASLEDEGYVE--VGKNAVLPSCSYTLPTSGTLVPMCMGKGFPCWSQCT 64  
 DB 3 LQVVLISLIHLADSVADVNDGVAAGISTILPCRY----NGAITSNMGHGTCSVFSCP 58  
 QY 65 NELIARTDERNVTYQSSRYQLKGLNKGDVSLIINKVTLDDHGYCCRIQPPGLNDKKL 124  
 DB 59 DGIWVTNGTHTYRKETRYKLLGNLSRRDVSLLTANTAVSDGIYCCRVKHSQWFNDMKI 118  
 QY 125 ELKLDIKAKAYT----- 136  
 DB 119 TISLKIIPPRTVPIVTRTSTVPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTL 178  
 QY 137 ----- 136  
 DB 179 LPTTTVPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTLP 238  
 QY 137 ----- 136  
 DB 239 TTTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPT 298  
 QY 137 -----PAQT-----AHGDTTASP-KITITERNGSETOVLTLNNNGTKIST 178  
 DB 299 PVATSPSPQPAETHPTVLGATRTQPTSSPLYSTTDSPTVSSDGLMNNQTOLOSP 358  
 QY 179 MADEIKDGETRT--AIHIGVSGAGLTLALIGVLLIKWYSCCKKKLSLITLNL 236  
 DB 359 -----EHSPOWNTTEGIVAGVCISV-LVLAIVGVIAKTYFF-KKEIQOLS-VSFSNH 410  
 QY 237 PPGIAGNAGAVRINSENIYTIENNV 263  
 DB 411 QFKTLQNAVKKEVHAEDNIY-IENNLX 436

RESULT 15  
 ID 046597 PRELIMINARY; PRT; 474 AA.  
 AC 046597;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE Hepatitis A virus cellular receptor 1 long form (Hepatitis A virus  
 cellular receptor 1 short form).  
 GN HAVCR-1.  
 OS Cercopithecus aethiops (Green monkey) (Griwet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxID=9534;  
 RN [1]  
 RP SEQUENCE FROM N. A.  
 RA Feigelestock D., Thompson P., Mattoo P., Kaplan G.G.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF043447; AAC39772.1; -;  
 DR EMBL; AF043446; AAC39771.1; -;  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG\_1.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 SQ SEQUENCE 474 AA; 51088 MM; 472DF987CA220524 CRC64;

Query Match 18.1%; Score 265.5; DB 6; Length 474;  
 Best Local Similarity 38.2%; Pred. No. 2.2e-14;  
 Matches 63; Conservative 24; Mismatches 71; Indels 7; Gaps 3;

QY 7 LNCVLLQLLLASLEDEGYVE--VGKNAVLPSCSYTLPTSGTLVPMCMGKGFPCWSQCT 64  
 DB 8 LQVVLISLIHLADSVADVNDGVAAGISTILPCRY----NGAITSNMGHGTCSVFSCP 63  
 QY 65 NELIARTDERNVTYQSSRYQLKGLNKGDVSLIINKVTLDDHGYCCRIQPPGLNDKKL 124  
 DB 64 DGIWVTNGTHTYRKETRYKLLGNLSRRDVSLLTANTAVSDGIYCCRVKHSQWFNDMKI 123  
 QY 125 ELKLDIKAKAY-TPAQTAHGDTTASPRTLITERNGSETOVLTL 168  
 DB 124 TISLKIIPPRTVPIVTRTSTVPTTTTLPTTTTLPTTTTLPTTTTLPTTTTL 168

Search completed: November 22, 2003, 05:43:21  
 Job time : 67.5584 secs

**THIS PAGE BLANK (USPTO)**



Query Match	100.0%	Score 1587	DB 17	Length 301
Best Local Similarity	100.0%	Pred. No. 8.8e-144		
Matches 301	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1 MFSLHPFCVLLILLILITRSESEVYRAVGONAYLPCTYPAAGNLVPCWKGACPV 60				
DB 1 MFSLHPFCVLLILLILITRSESEVYRAVGONAYLPCTYPAAGNLVPCWKGACPV 60				
QY 61 FEGGNVLRTERDENVYMSRYWLNQDFRKGDSVTIENVTLADSGIYCCRIQIPGINMD 120				
DB 61 FEGGNVLRTERDENVYMSRYWLNQDFRKGDSVTIENVTLADSGIYCCRIQIPGINMD 120				
QY 121 EKFNLKLVTKPAKVTAPLQDFPAAPPRLMTTGGHGAPEQTGLSPDINLTQISTLA 180				
DB 121 EKFNLKLVTKPAKVTAPLQDFPAAPPRLMTTGGHGAPEQTGLSPDINLTQISTLA 180				
QY 181 NELRSRLANDLRDGSATIRIGIYGAGICAGLALALFGLALIFKMSHSKEKIONLSLI 240				
DB 181 NELRSRLANDLRDGSATIRIGIYGAGICAGLALALFGLALIFKMSHSKEKIONLSLI 240				
QY 241 SLANLPSPGLANAVAEGRISEENITYIENVEVEPEPEYCYVSSROQSPQIGCREPM 300				
DB 241 SLANLPSPGLANAVAEGRISEENITYIENVEVEPEPEYCYVSSROQSPQIGCREPM 300				
QY 301 P 301				
DB 301 P 301				

RESULT 2  
AA97058  
ID AA97058 standard; Protein, 301 AA.  
AC  
AA97058;  
DT 31-OCT-2000 (first entry)  
DE Human T helper cell differentially expressed gene 200 product.  
KW T helper cell; differential expression; 200 gene; immunomodulator;  
KW anti-inflammatory; anti-arthritis; antibacterial; immunosuppressor;  
KW thymometric; anti-thyroid; anti-asthmatic; anti-allergic; antiviral;  
KW protozoacide; lymphocyte; modulator; gene therapy.  
XX

OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	/label= Signal_peptide
FT	Protein
FT	/label= Mature_protein
FT	Domain
FT	/label= Extracellular_domain
FT	Domain
FT	/label= Transmembrane_domain
FT	Domain
FT	/label= Cytoplasmic_domain
XX	
PN	US6084083-A.
XX	
PD	04-JUL-2000.
PF	
XX	28-MAR-1997; 97US-0829525.
XX	
PR	01-MAR-1996; 96US-0609583.
PR	03-MAR-1995; 95US-0398633.
XX	07-JUN-1995; 95US-0487748.
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
P1	Levinson DA;
XX	
DR	WPI; 2000-464385/40.
N-P8DB; AAA51904, AAA51905.	
PT	New isolated human 200 gene products or polypeptides, useful for treating and diagnosing immune disorders, especially T helper lymphocyte-related disorders
PS	Claim 1; Fig 24A-D; 107pp; English.
XX	
CC	Genes which are differentially expressed within and among T helper (TH) cells and TH cell subpopulations, e.g. TH0, TH1 and TH2 subpopulations, can be used diagnostically or as targets for therapeutic intervention.
CC	The polypeptides are useful for treating and diagnosing of immune disorders, especially T lymphocyte-related disorders. These disorders include chronic inflammatory diseases and disorders (e.g. Crohn's disease, reactive arthritis, Lyme disease, Hashimoto's thyroiditis or Graves' disease), or atopic conditions (e.g. asthma and allergy, including allergic rhinitis or food allergies). Also included are certain pathogen susceptibilities (e.g. leishmaniasis), and viral (e.g. HIV) or bacterial (e.g. tuberculosis or leptomatous leprosy) infections.
CC	
XX	
SQ	Sequence 301 AA;
Query Match	100.0%; Score 1587; DB 21; Length 301;
Best Local Similarity	100.0%; Pred. No. 8.8e-144;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MFSLPFDVLLLLLLLITRSSSEVEYRAVGQNAYLPFCFYTPAAPAGNLVPVCWGKACP 60
Db	1 MFSLHPFCVLLLLLLLLITRSSEVEYRAVEGQNNYLPCFYTAPAGNLVPCWGKACP 60
QY	61 FEGGNVLRTDERDVNVTSTRYMLNGDFRKDVSLTIENTVTLADSGIYCCKIQIPGINMD 120
Db	61 FEGGNVVLRTDERDVNVMTSTRYMLNGDFRKDVSLTIENTVTLADSGIYCCKRIQIPGINMD 120
QY	121 EKFNKLIVIRAKATPAPTLORDTAAFPRLMTTRHGAPARTOTLGSLPDINLTQISTLA 180
Db	121 EKFNKLIVIRAKATPAPTLORDTAAFPRLMTTRHGAPARTOTLGSLPDINLTQISTLA 180
QY	181 NELRDSRLANDLRDSGATIRIGIYICAGIALALIFGALLPKWTSHSXEKIQLNSLI 240
Db	181 NELRDSRLANDLRDSGATIRIGIYICAGIALALIFGALLPKWTSHSXEKIQLNSLI 240
QY	241 SLANLPBGLANAAYAEGRISENTYTIIEENVYEVEEPNEYCYVWSROQSPQIGCRPM 300



Db 241 SLANLPPSGLANAVAGIRSEENITYTEENYVEVEEPNEYCYVSSRQPSQPLGCRPAM 300  
Qy 301 P 301  
Db 301 P 301

RESULT 3  
AAM50223  
ID AAM50223 standard; Protein; 301 AA.  
XX  
AC AAM50223;  
XX  
DT 07-JAN-2002 (first entry)  
XX  
DE Human 200 gene product, preferentially expressed in TH1 cells.  
XX  
KW Human; 200 gene; T helper; T lymphocyte; T cell; TH1;  
KW receptor; differential expression; immune disorder; psoriasis;  
KW multiple sclerosis; insulin-dependent diabetes; antidiabetic;  
KW antiporiatic; diagnosis; therapy.  
XX

OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..20  
FT /label= Signal\_peptide  
FT Protein 21..301  
FT /label= Mature\_protein  
FT Domain 21..200  
FT /label= Extracellular\_domain  
FT Domain 201..224  
FT /label= Transmembrane\_domain  
FT Domain 225..301  
FT /label= Cytoplasmic\_domain  
XX  
PN US6288218-B1.  
XX  
PD 11-SEP-2001.  
XX  
PF 25-SEP-1997; 97US-09373399.  
XX  
PR 01-MAR-1996; 96US-0609583.  
PR 03-MAR-1995; 95US-0398633.  
PR 07-JUN-1995; 95US-0487748.  
XX  
PA (LEVI/) LEVINSON D A.  
XX  
PI Levinson DA;  
XX  
DR WPI; 2001-647189/74.  
DR N-PSDB; AAI70254; AAI70281.  
XX  
PT Detecting 200 gene expression in a sample, useful for treating and  
PT diagnosing immune disorders, especially T lymphocyte-related disorders,  
PT comprises detecting the presence of a 200 gene product or an RNA  
PT encoding the 200 gene product -  
XX  
PS Claim 1(a); Fig 24A-B; 108pp; English.  
XX  
XX The present sequence is that of the protein product of the human  
XX 200 gene (see AAI70254). It is a receptor of the Ig superfamily  
XX class. The 200 gene is preferentially expressed in mature, fully  
XX differentiated T helper subpopulation TH1 cells relative to  
XX subpopulation TH2 cells. The gene can be used diagnostically or as  
XX a target for therapeutic intervention for the treatment of immune  
XX disorders. A claimed method for diagnosing a TH cell subpopulation  
XX related immune disorder involves detecting the level of a 200 gene  
XX product, or an RNA encoding it, so that if the level differs from  
XX that in a control sample, the disorder is diagnosed. The gene  
XX product detected may be the present amino acid sequence, or  
XX amino acids 1-20, 1-200, 1-224, 30-128, 21-200, 21-301,  
XX 201-224, 201-301 or 224-301 of it. The immune disorder is

CC especially a TH1 cell subpopulation-related immune disorder, such  
CC as multiple sclerosis, psoriasis or insulin-dependent diabetes  
CC (claimed). In addition to the 200 gene, the invention provides  
CC other genes that are differentially expressed within and among  
CC TH cells and TH cell subpopulations and which can be used in  
CC methods for the diagnosis, prognosis, evaluation and treatment of  
CC TH cell subpopulation-related disorders, for the identification of  
CC subjects exhibiting a predisposition to such conditions, for  
CC monitoring patients undergoing clinical evaluation for the  
CC treatment of such disorders, and for monitoring the efficacy of  
CC compounds used in clinical trials. Other immune disorders that can  
CC be treated/diagnosed include Crohn's disease, reactive arthritis,  
CC Lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis,  
CC Grave's disease, contact dermatitis, graft rejection, graft versus  
CC host disease, sarcoidosis, atopic conditions, asthma, allergy,  
CC allergic rhinitis, food allergy, eosinophilia, conjunctivitis,  
CC glomerular nephritis, helminthic infection (e.g. leishmaniasis),  
CC viral infection (e.g. HIV), and bacterial infection (e.g.  
CC tuberculosis and lepromatous leprosy).  
XX  
SQ Sequence 301 AA;  
XX

Query Match 100.0%; Score 1587; DB 22; Length 301;  
Best Local Similarity 100.0%; Pred. No. 8, 8e-144;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSLPPDVCVLLLLLTSSSEVEYRAEQNNYLPCFTTPAPGMLVPCWKGACPV 60  
Db 1 MFSLPPDVCVLLLLLTSSSEVEYRAEQNNYLPCFTTPAPGMLVPCWKGACPV 60

Qy 61 FECSNVVLRDERDVNWTSRWYMLNGDFRKDVELTIENTVTLADSGIYCCRIGIPIGIMND 120  
Db 61 FECSNVVLRDERDVNWTSRWYMLNGDFRKDVELTIENTVTLADSGIYCCRIGIPIGIMND 120

Qy 121 EKENLKVIRKAVYTPAFTLQDFTPAAPPMLTTRGHPAETQTLGSLPDINLTQISTLA 180  
Db 121 EKENLKVIRKAVYTPAFTLQDFTPAAPPMLTTRGHPAETQTLGSLPDINLTQISTLA 180

Qy 181 NEFDSRLANDLRSGATIRIGIYIGAGICGALALIFGLIRKWSHSEKIQNSLI 240  
Db 181 NEFDSRLANDLRSGATIRIGIYIGAGICGALALIFGLIRKWSHSEKIQNSLI 240

Qy 241 SLANLPPSGLANAVAGIRSEENITYTEENYVEVEEPNEYCYVSSRQPSQPLGCRPAM 300  
Db 241 SLANLPPSGLANAVAGIRSEENITYTEENYVEVEEPNEYCYVSSRQPSQPLGCRPAM 300

Qy 301 P 301  
Db 301 P 301

RESULT 4  
AAU14409  
ID AAU14409 standard; Protein; 301 AA.  
XX  
AC AAU14409;  
XX  
DT 24-OCT-2001 (first entry)  
XX  
DE Human novel protein #280.  
XX  
XX Human; novel protein; antianemic; osteopathic; antiinflammatory;  
XX immunomodulatory; cytostatic; neuroprotective; vulnery; nootropic;  
XX anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
XX antibacterial; antiallergic; dermatological; haemostatic; antisthmatic;  
XX thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
KW tissue regeneration; immune disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200155437-A2.

Query	Match	Best Local Similarity	100.0%	Score 1587	DB 22	Length 301
Matches	301	Conservative	100.0%	Mismatches 0	Indels 0	Gaps 0
QY	1	MFSHLPDPCVLLILLILLITRSSEVYRAEVGONAVLPCTYTPAAAGNLVPCWKGACPV	60			
Db	1	MFSHLPDPCVLLILLILLITRSSEVYRAEVGONAVLPCTYTPAAAGNLVPCWKGACPV	60			
QY	61	FEQGNVLRTEREDVNVWTSRYWLNQDFPKGDSVLTIEENVTLADSGIYCRITQPIGIMND	120			
Db	61	FEQGNVLRTEREDVNVWTSRYWLNQDFPKGDSVLTIEENVTLADSGIYCRITQPIGIMND	120			
QY	121	EKENLKVTPAKKVPAPLQRPDPAAPFRMLTTTGHGGAETQTCGSPDINLTQISTLA	180			
Db	121	EKENLKVTPAKKVPAPLQRPDPAAPFRMLTTTGHGGAETQTCGSPDINLTQISTLA	180			
QY	181	NELRDSRLANDRDSGATIRIGIYAGICAGIALALIFGALIFPKWSSKKEIQNLSLI	240			
Db	181	NELRDSRLANDRDSGATIRIGIYAGICAGIALALIFGALIFPKWSSKKEIQNLSLI	240			
QY	241	SLANLPSSGLANAAVAGIRSEENIYTIIEENVYVEEPENYCYVSSRQOPSGPLGCRFAM	300			
Db	241	SLANLPSSGLANAAVAGIRSEENIYTIIEENVYVEEPENYCYVSSRQOPSGPLGCRFAM	300			
QY	301	P 301				
Db	301	P 301				

RESULT 5  
 AAB93838  
 ID AAB93838 standard; Protein: 301 AA.  
 XX AAB93838;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human protein sequence SEQ ID NO:13669.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EPI074617-AA.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 DR WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length CDNs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length CDNs -  
 XX  
 PS Claim 8; SEQ ID 13669; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length CDNs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length CDNs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length CDNs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 301 AA;  
 Query Match 100.0%; Score 1587; DB 22; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 8 8e-144;  
 Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MFSLPDCVLLLLLLLTSSSEVERAEGVGNAYVPCFTPAAPGNLVPCWKGKACPV 60

Db 1 MESHLPDCCVLLLLLLLITRSSEVYRAEVGQNAVLPCEYTPAAPGNLVPVCMGKACPV 60  
QY 61 FECCGNVVLRTDERDVNWTSRWYMLNGDFRKGDVSLTIENTVTLADSGIYCCRIQIPGIMND 120  
Db 61 FECCGNVVLRTDERDVNWTSRWYMLNGDFRKGDVSLTIENTVTLADSGIYCCRIQIPGIMND 120  
QY 121 EKFNKLVIRPAKVTPTLQORDFTAAFPRLMTTRGHGPAETQTLGSLPDINTLTQISTLA 180  
Db 121 EKFNKLVIRPAKVTPTLQORDFTAAFPRLMTTRGHGPAETQTLGSLPDINTLTQISTLA 180  
QY 181 NELRDSRLANDLRDSGATIRIGIYIGAGICAGLALALIFGALIFKWSHSEKIQNLSLI 240  
Db 181 NELRDSRLANDLRDSGATIRIGIYIGAGICAGLALALIFGALIFKWSHSEKIQNLSLI 240  
QY 241 SLANLPPSGLANAVAEGRSEENIYTIENYVEVEEPEEYCYVSSROQPSQPLGCRPFAM 300  
Db 241 SLANLPPSGLANAVAEGRSEENIYTIENYVEVEEPEEYCYVSSROQPSQPLGCRPFAM 300  
QY 301 P 301  
Db 301 P 301

RESULT 6  
AAB81518  
ID AAB81518 standard; Protein; 301 AA.  
XX AC AAB81518;  
XX 18-JUN-2001 (first entry)  
XX DE Human TH1 specific 200 gene product.  
XX KW Human; T helper cell; TH cell; TH1; TH2; immunomodulator;  
KW anti-inflammatory; antiallergic; dermatological; antiviral;  
KW antibacterial; T helper lymphocyte modulator; gene therapy;  
KW TH specific gene; 200 gene; immune disorder; inflammation;  
KW infection.  
XX OS Homo sapiens.  
XX PN US6204371-B1.  
XX PD 20-MAR-2001.  
XX PF 01-MAR-1996; 96US-0609583.  
XX PR 03-MAR-1995; 95US-0398633.  
XX PR 07-JUN-1995; 95US-0487748.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX P1 Levinson DA;  
XX DR WPI; 2001-272703/28.  
XX DR N-PSDB; AAF82616.  
XX FT New murine or human 200 genes and their corresponding polypeptides,  
PT useful for treating or diagnosing immune disorders, especially T helper  
PT lymphocyte-related disorders, e.g. inflammatory diseases or allergies  
XX  
XX PS Claim 5; Fig 24; 109pp; English.  
XX CC The present sequence is encoded by the human 200 gene, which is  
CC expressed at higher levels in T helper (TH)1 cells than in TH2 cells.  
CC The invention relates to an isolated nucleic acid molecule, which  
CC comprises the full length murine 200 gene or full length human 200  
CC gene nucleotide sequence. The nucleic acids are useful for treating  
CC or diagnosing immune disorders, especially T helper  
CC lymphocyte-related disorders, e.g. inflammatory diseases (e.g.  
CC Crohn's disease), multiple sclerosis, Grave's disease, contact  
CC dermatitis, psoriasis, asthma and allergies, or certain viral

CC (e.g. HIV) or bacterial (e.g. tuberculosis) infections.  
XX  
XX Sequence 301 AA;  
Query Match 100.0%; Score 1587; DB 22; Length 301;  
Best Local Similarity 100.0%; Pred. No. 8.8e-144;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MESHLPDCCVLLLLLLLITRSSEVYRAEVGQNAVLPCEYTPAAPGNLVPVCMGKACPV 60  
Db 1 MESHLPDCCVLLLLLLLITRSSEVYRAEVGQNAVLPCEYTPAAPGNLVPVCMGKACPV 60  
QY 61 FECCGNVVLRTDERDVNWTSRWYMLNGDFRKGDVSLTIENTVTLADSGIYCCRIQIPGIMND 120  
Db 61 FECCGNVVLRTDERDVNWTSRWYMLNGDFRKGDVSLTIENTVTLADSGIYCCRIQIPGIMND 120  
QY 121 EKFNKLVIRPAKVTPTLQORDFTAAFPRLMTTRGHGPAETQTLGSLPDINTLTQISTLA 180  
Db 121 EKFNKLVIRPAKVTPTLQORDFTAAFPRLMTTRGHGPAETQTLGSLPDINTLTQISTLA 180  
QY 181 NELRDSRLANDLRDSGATIRIGIYIGAGICAGLALALIFGALIFKWSHSEKIQNLSLI 240  
Db 181 NELRDSRLANDLRDSGATIRIGIYIGAGICAGLALALIFGALIFKWSHSEKIQNLSLI 240  
QY 241 SLANLPPSGLANAVAEGRSEENIYTIENYVEVEEPEEYCYVSSROQPSQPLGCRPFAM 300  
Db 241 SLANLPPSGLANAVAEGRSEENIYTIENYVEVEEPEEYCYVSSROQPSQPLGCRPFAM 300  
QY 301 P 301  
Db 301 P 301

RESULT 7  
AAB51104  
ID AAB51104 standard; Protein; 301 AA.  
XX AC AAB51104;  
XX 20-MAR-2001 (first entry)  
XX DE Human 200 gene product amino acid sequence SEQ ID NO:24.  
XX KW Treatment; diagnosis; immune disorder; mast cell related disorder;  
KW T-helper lymphocyte-related disorder; ischemic disorder;  
KW identification; vasodilator; cardiac; antianginal; angina pectoris;  
KW ischaemic renal disease; myocardial ischaemia; myocardial infarction;  
KW cortical infarction; ischaemic injury; kidney transplant.  
XX OS Homo sapiens.  
XX PN WO200073498-A1.  
XX PD 07-DEC-2000.  
XX PF 31-MAY-2000; 2000WO-US14986.  
XX PR 02-JUN-1999; 99US-0324986.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX P1 Levinson DA, Lloyd CM, McCarthy SA;  
XX DR WPI; 2001-016510/02.  
XX DR N-PSDB; AAC90993.  
XX FT Ameliorating a symptom of an ischemic disorder or injury in a mammal  
PT e.g. ischemic renal disease or myocardial ischemia, by administering a  
PT 200 gene product (S1), a nucleic acid encoding (S1) or an antibody  
PT directed against (S1) -  
XX  
XX PS Claim 10; Fig 24; 309pp; English.  
XX CC



PN US641417-B1.  
 XX  
 PD 02-JUL-2002.  
 XX  
 PF 12-MAY-1999; 99US-0310367.  
 XX  
 PR 01-MAR-1996; 96US-0609583.  
 PR 28-MAR-1997; 97US-0829525.  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Levinson DA;  
 XX  
 DR MPI; 2002-641576/69.  
 DR N-PSDB; ABS53322.  
 XX  
 PT Novel polypeptides which are differentially expressed within and among  
 PT T helper cells and cell populations, useful for treating immune  
 PT disorders, especially T helper cell subpopulation-related disorders -  
 XX  
 PS Example; Fig 24; 109pp; English.  
 XX  
 CC This invention relates to a novel isolated polypeptide which is a T  
 CC helper (Th1) cell subpopulation specific gene product. The invention  
 CC also discloses other Th1 or Th2 specific gene products. The proteins of  
 CC the invention may have antiinflammatory, antidiabetic, antithyroid,  
 CC antiascitic, antiallergic, virocidic, antiautistic, dermatological,  
 CC antiproliferative, nephrotropic and immunosuppressive activities and may be  
 CC used as a regulator of the immune response. The proteins of the  
 CC invention are also useful to reduce the level of Th2 cell activity for  
 CC treating Th1 cell subpopulation-related disorders including atopic  
 CC conditions such as asthma and allergy including allergic rhinitis, the  
 CC effects of pathogen, including viral infection, chronic inflammatory  
 CC diseases such as Crohn's disease, arthritis, diabetes, thyroiditis,  
 CC dermatitis, psoriasis, glomerular nephritis, organ-specific  
 CC autoimmunity, graft rejection and graft versus host disease. The  
 CC present sequence represents the human 200 gene protein of the invention.  
 XX  
 SQ Sequence 301 AA;  
 Query Match 100.0%; Score 1587; DB 23; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-144;  
 Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MFSLPFCVLLLLLTSSSEVYRAVGNAYLPCTYPAAGNIVPCWKGACPV 60  
 DB 1 MFSLPFCVLLLLLTSSSEVYRAVGNAYLPCTYPAAGNIVPCWKGACPV 60  
 QY 61 FECGNVLRTERDVNWTSRWYLNQDPRKGVSLTIENVTLADSGIYCCRIQIPGINND 120  
 DB 61 FECGNVLRTERDVNWTSRWYLNQDPRKGVSLTIENVTLADSGIYCCRIQIPGINND 120  
 QY 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 DB 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 QY 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 DB 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 QY 181 NELRPSRLANDIRDSGATIRIGIYAGICAGLALALIFGALIFKWSHSEKIONLSLI 240  
 DB 181 NELRPSRLANDIRDSGATIRIGIYAGICAGLALALIFGALIFKWSHSEKIONLSLI 240  
 QY 241 SLANLPPSGLANAVAGIRSEENITTEENVYVEPEPEYCYVSSROOPSGPLGCRPAM 300  
 DB 241 SLANLPPSGLANAVAGIRSEENITTEENVYVEPEPEYCYVSSROOPSGPLGCRPAM 300  
 QY 301 P 301  
 DB 301 P 301  
 RESULT 10  
 ABP70445

ID ABP70445 standard; Protein; 301 AA.  
 XX  
 AC ABP70445;  
 XX  
 DT 22-APR-2003 (first entry)  
 XX  
 DE Amino acid sequence of human TIM-3 allele 2.  
 XX  
 DE T cell immunoglobulin domain; mucin domain; TIM-1; TIM-2; TIM-3;  
 XX TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;  
 XX myelodysplastic syndrome; airway hyperactivity; cancer; asthma;  
 XX allergic T cell response; autoimmune disease.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 140  
 FT /note="encoded by CCG"  
 XX  
 PN MO200302722-A2.  
 XX  
 PD 09-JAN-2003.  
 XX  
 PF 01-JUL-2002; 2002WO-US20890.  
 XX  
 PR 29-JUN-2001; 2001US-302344P.  
 XX  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 XX  
 PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;  
 DR MPI; 2003-210268/20.  
 DR N-PSDB; ABZ68339.  
 XX  
 PT New nucleic acid comprising a mammalian T cell immunoglobulin domain  
 PT and Mucin domain gene sequences, useful for treating cancer or asthma,  
 PT allergy, eczema or autoimmune disease -  
 XX  
 PS Claim 10; Page 90; 94pp; English.  
 XX  
 CC The present sequence is a human T cell immunoglobulin domain and  
 CC mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1,  
 CC TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with  
 CC conserved IGV and mucin domains. The locus comprising the TIM family is  
 CC genetically associated with immune dysfunction, including asthma. The  
 CC TIM gene family is located within a region of human chromosome 5 that  
 CC is commonly deleted in malignancies and myelodysplastic syndrome.  
 CC Variants of TIM-1 and TIM-3 are associated with susceptibility to  
 CC airway hyperactivity and allergic T cell responses, and other variants  
 CC associated with protection against these responses. T cells express TIM  
 CC proteins, which critically regulate CD4 T cell differentiation. Th1  
 CC cells preferentially express TIM-3, while Th2 cells preferentially  
 CC express TIM-1. TIM polypeptides and polynucleotides are useful for  
 CC treating cancer, asthma, allergies, eczema or autoimmune diseases.  
 XX  
 SQ Sequence 301 AA;  
 Query Match 100.0%; Score 1587; DB 24; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-144;  
 Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MFSLPFCVLLLLLTSSSEVYRAVGNAYLPCTYPAAGNIVPCWKGACPV 60  
 DB 1 MFSLPFCVLLLLLTSSSEVYRAVGNAYLPCTYPAAGNIVPCWKGACPV 60  
 QY 61 FECGNVLRTERDVNWTSRWYLNQDPRKGVSLTIENVTLADSGIYCCRIQIPGINND 120  
 DB 61 FECGNVLRTERDVNWTSRWYLNQDPRKGVSLTIENVTLADSGIYCCRIQIPGINND 120  
 QY 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 DB 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 QY 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 DB 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180

QY 181 NEI.RD.SRLAND.RD.SGATIRIGIYAGICAGLALALIFGALIFKMYSHSKKEIQLNLSLI 240  
 CC |||||||  
 CC antiinflammatory, antiarthritic, antidiabetic, neuroprotective,  
 CC dermatological, antihypertensive, antipruritic, antineoplastic,  
 CC and antiallergic activity and can act as CD4, CD4 and interleukin  
 CC agonists. The methods and compositions of the present invention are  
 CC useful for the identification and therapeutic use of compounds as  
 CC treatments of helminthic, bacterial and viral infections and immune  
 CC disorders such as Crohn's disease, reactive arthritis, diabetes,  
 CC multiple sclerosis, Hashimoto's thyroiditis, Grave's disease, contact  
 CC dermatitis, psoriasis, graft rejection, graft versus host disease,  
 CC asthma, allergy and glomerulonephritis. They can also be used for gene  
 CC therapy, for the diagnostic evaluation and prognosis of TH cell  
 CC subpopulation-related disorders, identification of subjects exhibiting a  
 CC predisposition to such conditions, monitoring undergoing clinical  
 CC evaluation and efficacy for the treatment of the disorders. This sequence  
 CC represents a T-Helper cell associated polypeptide described in the  
 CC disclosure of the invention.

DB 241 SLANLPPSGLANAVAGIRSEENIYTIENVEVEEPNEYYCVSSROQPSOPLGCRFAM 300  
 CC |||||||  
 CC 241 SLANLPPSGLANAVAGIRSEENIYTIENVEVEEPNEYYCVSSROQPSOPLGCRFAM 300

QY 301 P 301  
 CC |  
 CC 301 P 301

DB 301 P 301

RESULT 11  
 ABG73718  
 ID ABG73718 standard; Protein; 301 AA.  
 AC ABG73718;  
 XX  
 XX  
 DT 09-APR-2003 (first entry)  
 XX  
 DE Human TH1-associated 200 protein SEQ ID 24.  
 KM T-cell receptor; TH; T helper cell; 103 gene; TH2 cell; human;  
 KM TH2 cell marker; TH1 cell; protozoacide; antibacterial; virucide;  
 KM immunosuppressive; antiinflammatory; antipruritic; antidiabetic;  
 KM neuroprotective; dermatological; antihypertensive; antipruritic; helminthic;  
 KM nephrotoxic; antiallergic; antiallergic; CD4 agonist; CD4 agonist;  
 KM interleukin agonist; bacterial; viral infection; immune disorder;  
 KM Crohn's disease; reactive arthritis; diabetes; multiple sclerosis;  
 KM Hashimoto's thyroiditis; Grave's disease; contact dermatitis; psoriasis;  
 KM graft rejection; graft versus host disease; asthma; glomerulonephritis;  
 KM allergy; gene therapy; TH cell subpopulation.

XX Homo sapiens.  
 XX OS  
 PN US6455685-B1.  
 PD 24-SEP-2002.  
 XX  
 XX 27-FEB-1998; 98US-0032337.  
 PF  
 XX 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 PR 01-MAR-1996; 96US-0609583.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA  
 PI Levinson DA;  
 XX  
 XX WPI, 2003-066247/06.  
 DR N-PSDB; ABQ77040, ABQ77041.  
 DR  
 XX  
 PT Identifying a test compound that binds to a 103 gene product or is a  
 PT TH2 cell marker, useful for the identification and therapeutic use of  
 PT compounds as treatments of helminthic, bacterial and viral infections,  
 PT and immune disorders -  
 XX  
 XX Example 9; Figure 24A-C; 131p; English.

PS This invention describes a novel method for identifying a test compound  
 CC that binds to a 103 gene product or is a TH2 cell marker. The method  
 CC comprises contacting a test compound with an immobilised 103 gene  
 CC product, removing unbound test compound or separating the complex from  
 CC the reaction mixture, and detecting the complex. Identifying a test  
 CC compound that binds to a 103 gene product alternatively comprises: (a)  
 CC contacting a test compound with a cell engineered to express a 103 gene  
 CC product or co-expressing a 103 gene product and a test compound in a  
 CC cell; and (b) removing unbound 103 gene product. Identifying a test  
 CC compound that is a TH2 cell marker further comprises detecting a complex,  
 CC where contacting the test compound with a TH2 and TH1 cell for the test  
 CC compound to bind either cell and detecting binding where it indicates the  
 CC test compound as a TH2 cell marker. The products described in the

CC invention have protozoacide, antibacterial, virucide, immunosuppressive,  
 CC antiinflammatory, antiarthritic, antidiabetic, neuroprotective,  
 CC dermatological, antihypertensive, antipruritic, antineoplastic,  
 CC and antiallergic activity and can act as CD4, CD4 and interleukin  
 CC agonists. The methods and compositions of the present invention are  
 CC useful for the identification and therapeutic use of compounds as  
 CC treatments of helminthic, bacterial and viral infections and immune  
 CC disorders such as Crohn's disease, reactive arthritis, diabetes,  
 CC multiple sclerosis, Hashimoto's thyroiditis, Grave's disease, contact  
 CC dermatitis, psoriasis, graft rejection, graft versus host disease,  
 CC asthma, allergy and glomerulonephritis. They can also be used for gene  
 CC therapy, for the diagnostic evaluation and prognosis of TH cell  
 CC subpopulation-related disorders, identification of subjects exhibiting a  
 CC predisposition to such conditions, monitoring undergoing clinical  
 CC evaluation and efficacy for the treatment of the disorders. This sequence  
 CC represents a T-Helper cell associated polypeptide described in the  
 CC disclosure of the invention.

QY 1 MESH.PPDCVLLLLLLLRSEVEYRAVGONAVLPCYTPAAGNLVPCMGKACPV 60  
 CC |||||||  
 CC 1 MESH.PPDCVLLLLLLLRSEVEYRAVGONAVLPCYTPAAGNLVPCMGKACPV 60

DB 61 PEGGVVLTEDRDVNYWTSRYWLNQDFPKGVSJTIENVTADSGIYCCR1QIPQIMND 120  
 CC |||||||  
 CC 61 PEGGVVLTEDRDVNYWTSRYWLNQDFPKGVSJTIENVTADSGIYCCR1QIPQIMND 120

QY 61 PEGGVVLTEDRDVNYWTSRYWLNQDFPKGVSJTIENVTADSGIYCCR1QIPQIMND 120  
 CC |||||||  
 CC 61 PEGGVVLTEDRDVNYWTSRYWLNQDFPKGVSJTIENVTADSGIYCCR1QIPQIMND 120

DB 121 EKFNKLVIYKPAKVPAPLQDFTAAPRMLTTGHPAETQIGSLPDINLTQISTLA 180  
 CC |||||||  
 CC 121 EKFNKLVIYKPAKVPAPLQDFTAAPRMLTTGHPAETQIGSLPDINLTQISTLA 180

QY 121 EKFNKLVIYKPAKVPAPLQDFTAAPRMLTTGHPAETQIGSLPDINLTQISTLA 180  
 CC |||||||  
 CC 121 EKFNKLVIYKPAKVPAPLQDFTAAPRMLTTGHPAETQIGSLPDINLTQISTLA 180

DB 181 NEI.RD.SRLAND.RD.SGATIRIGIYAGICAGLALALIFGALIFKMYSHSKKEIQLNLSLI 240  
 CC |||||||  
 CC 181 NEI.RD.SRLAND.RD.SGATIRIGIYAGICAGLALALIFGALIFKMYSHSKKEIQLNLSLI 240

QY 241 SLANLPPSGLANAVAGIRSEENIYTIENVEVEEPNEYYCVSSROQPSOPLGCRFAM 300  
 CC |||||||  
 CC 241 SLANLPPSGLANAVAGIRSEENIYTIENVEVEEPNEYYCVSSROQPSOPLGCRFAM 300

DB 241 SLANLPPSGLANAVAGIRSEENIYTIENVEVEEPNEYYCVSSROQPSOPLGCRFAM 300  
 CC |||||||  
 CC 241 SLANLPPSGLANAVAGIRSEENIYTIENVEVEEPNEYYCVSSROQPSOPLGCRFAM 300

QY 301 P 301  
 CC |  
 CC 301 P 301

DB 301 P 301

RESULT 12  
 ABP70444  
 ID ABP70444 standard; Protein; 301 AA.  
 AC ABP70444;  
 XX  
 XX  
 DT 22-APR-2003 (first entry)  
 XX  
 DE Amino acid sequence of human TIM-3 allele 1.  
 DE  
 XX T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;  
 XX TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;  
 XX myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;  
 XX allergic T cell response; autoimmune disease.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 140  
 FT /note="encoded by CDS"  
 PN W02003002722-A2.  
 XX  
 XX 09-JAN-2003.

XX 01-JUN-2002; 2002WO-US20890.  
 PF  
 XX 29-JUN-2001; 2001US-302344P.  
 PR  
 XX (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PA  
 XX McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;  
 PI  
 XX WPI; 2003-210268/20.  
 DR  
 XX N-PSDB; ABZ68338.  
 DR  
 XX New nucleic acid comprising a mammalian T cell immunoglobulin domain  
 PT and mucin domain gene sequences, useful for treating cancer or asthma,  
 PT allergy, eczema or autoimmune disease -  
 XX  
 XX Claim 10; Page 89; 94pp; English.  
 XX  
 XX The present sequence is a human T cell immunoglobulin domain and  
 CC mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1,  
 CC TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with  
 CC conserved IgV and mucin domains. The locus comprising the TIM family is  
 CC genetically associated with immune dysfunction, including asthma. The  
 CC TIM gene family is located within a region of human chromosome 5 that  
 CC is commonly deleted in malignancies and myelodysplastic syndrome.  
 CC Variants of TIM-1 and TIM-3 are associated with susceptibility to  
 CC airway hyperreactivity and allergic T cell responses and other variants  
 CC associated with protection against these responses. T cells express TIM  
 CC proteins, which critically regulate CD4 T cell differentiation. Th1  
 CC cells preferentially express TIM-3, while Th2 cells preferentially  
 CC express TIM-1. TIM polypeptides and polynucleotides are useful for  
 CC treating cancer, asthma, allergies, eczema or autoimmune diseases.  
 XX  
 XX Sequence 301 AA:  
 SQ  
 Query Match 99.6%; Score 1581; DB 24; Length 301;  
 Best Local Similarity 99.7%; Pred. No. 3.3e-143;  
 Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MFSLPFCVLLLLLLLTSSSEVYRAEYVQGNALPCFYPPAAGNLVPCWKGACPV 60  
 Db 1 MFSLPFCVLLLLLLLTSSSEVYRAEYVQGNALPCFYPPAAGNLVPCWKGACPV 60  
 QY 61 FECGNVLTARDERDVNYWTSRYWLNQDPRKGDVSLITENVTLADSGIYCCRIQIPGINND 120  
 Db 61 FECGNVLTARDERDVNYWTSRYWLNQDPRKGDVSLITENVTLADSGIYCCRIQIPGINND 120  
 QY 121 EKFNKLVIKPAKVTAPATLQRFDAFPRMLTTGHPAETQTGLSPDINTLQISTLA 180  
 Db 121 EKFNKLVIKPAKVTAPATLQRFDAFPRMLTTGHPAETQTGLSPDINTLQISTLA 180  
 QY 181 NELDRSLANDLRDSGATIRIGIYAGICAGLALALFGALIFWYSHSKKIONLSLI 240  
 Db 181 NELDRSLANDLRDSGATIRIGIYAGICAGLALALFGALIFWYSHSKKIONLSLI 240  
 QY 241 SLANLPPSGLANNAVEGIRSEENITYIENYVEVEPEPNEYCYVSSROOPSLGCRFAM 300  
 Db 241 SLANLPPSGLANNAVEGIRSEENITYIENYVEVEPEPNEYCYVSSROOPSLGCRFAM 300  
 QY 301 P 301  
 Db 301 P 301  
 RESULT 13  
 AAU14173  
 ID AAU14173 standard; Protein; 301 AA.  
 XX AAU14173;  
 AC  
 XX 24-OCT-2001 (first entry)  
 DT  
 XX Human novel protein #44.  
 DE

XX Human; novel protein; cytostatic; neuroprotective; vulnerrary; nootropic;  
 KW immunomodulatory; antianemic; osteoporotic; anti-inflammatory;  
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antiaesthetic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200155437-A2.  
 PN  
 XX 02-AUG-2001.  
 PD  
 XX 25-JAN-2001; 2001WO-US02623.  
 PF  
 XX 25-JAN-2000; 2000US-0491404.  
 PR  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Tang YT, Liu C, Dzmanac RT;  
 PI  
 XX WPI; 2001-451939/48.  
 DR  
 XX N-PSDB; AAS22478.  
 DR  
 XX Isolated polypeptides useful for treating anti-inflammatory diseases,  
 PT nervous system disorders, and for regenerating bone and cartilage -  
 XX  
 XX Example 4; Page 550; 894pp; English.  
 XX  
 XX The invention relates to polynucleotides encoding novel human  
 CC proteins or their active domains. The polypeptides, polynucleotides and  
 CC antibodies raised against the polypeptides are used in a method of  
 CC treatment of a mammal and prevention of disorders caused by the aberrant  
 CC protein expression or activity. The polypeptides can be used as  
 CC molecular weight markers, food supplements, and in antibody production.  
 CC The polypeptides are used to identify compounds which bind to the  
 CC polypeptides. Polynucleotides of the invention are used as probes and  
 CC primers, for sequencing, for chromosome or gene mapping, in the  
 CC production of recombinant proteins, and in generating anti-sense DNA or  
 CC RNA and in gene therapy. Polypeptides of the invention can be used to  
 CC target drugs to a tumour, in assays to determine biological activity, to  
 CC raise antibodies/elicits an immune response, to determine quantitative  
 CC protein levels, as tissue markers, and to isolate receptors or ligands.  
 CC Polypeptides of the invention may also be useful in creating placental  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,  
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,  
 CC anti-inflammatory diseases, nervous system disorders, and infection.  
 CC The present sequence represents a protein of the invention.  
 XX  
 XX Sequence 301 AA:  
 SQ  
 Query Match 99.6%; Score 1580; DB 22; Length 301;  
 Best Local Similarity 99.7%; Pred. No. 4.1e-143;  
 Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MFSLPFCVLLLLLLLTSSSEVYRAEYVQGNALPCFYPPAAGNLVPCWKGACPV 60  
 Db 1 MFSLPFCVLLLLLLLTSSSEVYRAEYVQGNALPCFYPPAAGNLVPCWKGACPV 60  
 QY 61 FECGNVLTARDERDVNYWTSRYWLNQDPRKGDVSLITENVTLADSGIYCCRIQIPGINND 120  
 Db 61 FECGNVLTARDERDVNYWTSRYWLNQDPRKGDVSLITENVTLADSGIYCCRIQIPGINND 120  
 QY 121 EKFNKLVIKPAKVTAPATLQRFDAFPRMLTTGHPAETQTGLSPDINTLQISTLA 180  
 Db 121 EKFNKLVIKPAKVTAPATLQRFDAFPRMLTTGHPAETQTGLSPDINTLQISTLA 180

Db 121 EKFNKLKLVKPAKVPAPTLQGRDFTAAFPRLMTTRGHGPAETQTLGSLPDINLTQISTLA 180  
QY 181 NEIRDRLANDLRDSCATTIRIGIYTGAGTCAGALALIFGALLFKWYSHSKKXIONLSLI 240  
Db 181 NEIRDRLANDLRDSCATTIRIGIYTGAGTCAGALALIFGALLFKWYSHSKKXIONLSLI 240  
QY 241 SLANLPPSGLANAVAGIRSEENITYTEENVYEEVEEPNEYCYVSSROQPSOPLGCRFAM 300  
Db 241 SLANLPPSGLANAVAGIRSEENITYTEENVYEEVEEPNEYCYVSSROQPSOPLGCRFAM 300  
QY 301 P 301  
Db 301 P 301  
RESULT 14  
ID ABB90396 standard; Protein; 301 AA.  
AC ABB90396;  
DT 24-MAY-2002 (first entry)  
XX Human polypeptide SEQ ID NO 2772.  
DE  
XX Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;  
KW antidiabetic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein.  
XX  
OS Homo sapiens.  
PN WO200190304-A2.  
XX 29-NOV-2001.  
PD 18-MAY-2001; 2001WO-US16450.  
XX 19-MAY-2000; 2000US-205515P.  
PR (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Birse CE, Rosen CA;  
PI WPI, 2002-122018/16.  
DR N-PSDB; ABL90805.  
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
PT prevention of neural, immune system, muscular, reproductive,  
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
PT disorders -  
XX  
PS Claim 11; SEQ ID NO 2772; 2081pp + Sequence Listing; English.  
XX  
CC The invention relates to novel genes (ABL9449-ABL90853) and proteins  
CC (ABB90404-ABB90444) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WFO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 301 AA;  
Query Match 99.6%; Score 1580; DB 23; Length 301;  
Best Local Similarity 99.7%; Pred. No. 4, 1e-143;  
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MFSHLPEFDCVLLILLILLTRSSSEVEYRAVGONALPCYTPAAGNLVPCWGKACAV 60  
Db 1 MFSHLPEFDCVLLILLILLTRSSSEVEYRAVGONALPCYTPAAGNLVPCWGKACAV 60  
QY 61 FECGNVLTDRDQVNYWTSRYWLNQDFPKGDSLTIEVTLADSGIYCCRIQIPIIND 120  
Db 61 FECGNVLTDRDQVNYWTSRYWLNQDFPKGDSLTIEVTLADSGIYCCRIQIPIIND 120  
QY 121 EKFNKLKLVKPAKVPAPTLQGRDFTAAFPRLMTTRGHGPAETQTLGSLPDINLTQISTLA 180  
Db 121 EKFNKLKLVKPAKVPAPTLQGRDFTAAFPRLMTTRGHGPAETQTLGSLPDINLTQISTLA 180  
QY 181 NEIRDRLANDLRDSCATTIRIGIYTGAGTCAGALALIFGALLFKWYSHSKKXIONLSLI 240  
Db 181 NEIRDRLANDLRDSCATTIRIGIYTGAGTCAGALALIFGALLFKWYSHSKKXIONLSLI 240  
QY 241 SLANLPPSGLANAVAGIRSEENITYTEENVYEEVEEPNEYCYVSSROQPSOPLGCRFAM 300  
Db 241 SLANLPPSGLANAVAGIRSEENITYTEENVYEEVEEPNEYCYVSSROQPSOPLGCRFAM 300  
QY 301 P 301  
Db 301 P 301  
RESULT 15  
ID ABB90396 standard; Protein; 401 AA.  
AC ABB90396;  
DT 22-MAY-2003 (first entry)  
XX  
XX NOVX protein sequence SEQ ID NO 77.  
DE  
XX Hepatotropic; immunosuppressive; cardiant; hypertensive; tranquilizer;  
KW vulnerary; virucide; antibacterial; protozoacide; fungicide; nocotropic;  
KW antiparasitic; neuroprotective; cerebroprotective; antiparkinsonian;  
KW anticonvulsant; antidiabetic; analgesic; dermatological; keratolytic;  
KW antiseborrheic; antihemetic; antidiarrhetic; antiinflammatory; anti-HIV;  
KW cytostatic; antiaesthetic; antipsoriatic; hypotensive; osteopathic;  
KW antitumor; anorectic; antidiabetic; antiallergic; haemostatic;  
KW neuroleptic; antidepressant; antifertility; NOVX, human disease;  
KW NOVX-associated disorder; trauma; viral; bacterial; fungal; protozoal;  
KW parasitic infection; Alzheimer's disease; stroke; forensic biology;  
KW immunogen; non-human transgenic animal; gene therapy.  
XX  
OS Unidentified.  
PN WO200281517-A2.  
XX 17-OCT-2002.  
PD 22-JAN-2002; 2002WO-US02064.  
XX 19-JAN-2001; 2001US-262892P.  
PR 23-JAN-2001; 2001US-263598P.  
PR 24-JAN-2001; 2001US-263799P.  
PR 25-JAN-2001; 2001US-264117P.  
PR 25-JAN-2001; 2001US-264139P.  
PR 26-JAN-2001; 2001US-264478P.  
PR 30-JAN-2001; 2001US-263511P.  
PR 02-MAR-2001; 2001US-272870P.  
PR 14-MAR-2001; 2001US-275927P.  
PR 14-MAR-2001; 2001US-275927P.  
PR 15-MAR-2001; 2001US-276449P.



20-MAR-2001; 2001US-277358P.  
 23-MAR-2001; 2001US-278151P.  
 29-MAR-2001; 2001US-279857P.  
 20-APR-2001; 2001US-285140P.  
 20-APR-2001; 2001US-285141P.  
 30-APR-2001; 2001US-287484P.  
 17-MAY-2001; 2001US-291701P.  
 08-JUN-2001; 2001US-296960P.  
 10-JUL-2001; 2001US-304353P.  
 10-JUL-2001; 2001US-304355P.  
 12-JUL-2001; 2001US-304886P.  
 09-AUG-2001; 2001US-311289P.  
 13-AUG-2001; 2001US-311975P.  
 16-AUG-2001; 2001US-312937P.  
 18-OCT-2001; 2001US-330227P.  
 29-NOV-2001; 2001US-334198P.

(CURA-) CURAGEN CORP.

Decristofaro MF, Padigaru M, Miller C, Tchernev V, Zhong H;  
 Zhong M, Anderson D, Ballinger R, Gerlach V, Spylek KA;  
 Rastelli L, Kehuda R, Guo X, Zernhusen B, Andrew D, Mezes P;  
 Patcuajan M, Burgess CE, Eisen A, Wolenc A, Baumgartner U;  
 Shumke RA, Gusev V, Vernet CM, Taupier RJ, Pena C, Shenoy S;  
 Li L, Casman S, Bolzog F, Fernandes E, Smithson G, Malyankar U;  
 Tallon B, Liu X;

WPI: 2003-058504/05.  
 N-PSDB: ABT33381.

New polypeptides, designated as NOVX, useful for diagnosing and  
 treating infections, neurological diseases, cancer, allergy, and bone,  
 immunological, skin, renal, brain, muscle and autoimmune disorders -

Claim 1; Page 209; 672pp; English.

The invention relates to a novel isolated polypeptide, designated NOVX  
 (NOV1 - 33), consisting of a mature form of one of 61 sequences, given  
 in the specification, or its variant, where amino acid residue(s) in the  
 variant differ from the mature form, provided that the variant differs  
 in not more than 15 % of the amino acids from the sequence of the mature  
 form. The NOVX polypeptides, nucleic acids encoding the polypeptides, and  
 an antibody to the polypeptides, are useful for treating or preventing a  
 NOVX-associated disorder in humans and for treating a syndrome associated  
 with a human disease (NOVX-associated disorder). NOVX polypeptides and  
 the encoding nucleic acids, are useful for determining the presence of or  
 predisposition to a disease associated with altered levels of NOVX  
 polypeptide and polynucleotide, by measuring the level of polypeptide  
 expression or the amount of nucleic acid from a mammal and comparing it  
 with another mammal not having or not predisposed to the disease. NOVX  
 polypeptide is also useful for identifying an agent that binds to NOVX  
 and a cell expressing NOVX is useful for identifying an agent that  
 modulates the expression or activity of NOVX. The antibodies and a  
 polypeptide having 95 % sequence identity to NOVX polypeptide are useful  
 for treating a pathological state in a mammal. The antibodies are also  
 useful for determining the presence or amount of NOVX in a sample. NOVX  
 polypeptides, polynucleotides and antibodies specific for the  
 polypeptides are useful for treating or preventing disorders or syndromes  
 including trauma, viral, bacterial, fungal, protozoal, and parasitic  
 infections. They can also treat disorders such as e.g., Alzheimer's  
 disease or a stroke. The NOVX encoding nucleic acids are useful for  
 expressing the NOVX proteins, to detect NOVX mRNA, or a genetic lesion in  
 a NOVX gene and to modulate NOVX activity. NOVX sequences are also useful  
 for identifying a cell or tissue type in a biological sample, to amplify  
 DNA sequences from very small biological samples such as tissues e.g.  
 hair or skin or body fluids in forensic biology and as primers and probes  
 for use in identifying and/or cloning NOVX homologues in other cell  
 types. The NOVX proteins are useful as an immunogen to generate  
 antibodies which are useful for diagnostically monitoring protein levels  
 and modulating NOVX activity. Cells comprising NOVX nucleic acids are  
 useful for producing non-human transgenic animals which are useful for  
 studying the function and/or activity of NOVX protein and for identifying  
 and/or evaluating modulators of NOVX protein activity. The NOVX nucleic

CC acids can be used in gene therapy. This sequence represents a NOVX  
 CC protein of the invention.

XX Sequence 401 AA;

Query Match 91.3%; Score 1449; DB 24; Length 401;

Best Local Similarity 72.3%; Pred. No. 2,4e-130;

Matches 290; Conservative 3; Mismatches 8; Indels 100; Gaps 2;

1 MFSLPPDCVLLLLLL-----  
 1 MFSLPPDCVLLLLLTTLPSRVHRIGPSPRAVFLKVRONGVKQSDNAFVSGGD 60  
 18 -----LTR-----  
 61 RTGDELSTICLAQUTRFLPVKIMGSSFLFILDMACSPFEMELPHSPLBSWLSFLYAG 120  
 21 SSEVEYRAEVQNNAYLPCFYTPAAPGNLVPVCWKGACVPEECGVVLTDERDVTWTS 80  
 121 SSEVEYRAEVQNNAYLPCFYTPAAPGNLVPVCWKGACVPEECGVVLTDERDVTWTS 180  
 81 RYWLNGPFRGQVSLTENTVTLADSGIYCCRIQIRGIMDEKFKLVIRAKVTPA 140  
 181 RYWLNGPFRGQVSLTENTVTLADSGIYCCRIQIRGIMDEKFKLVIRAKVTPA 240  
 141 QRDFTAFPRMLTTRGHGPAETOTLGSIPDINLTQISTLANELRDSRLANDRDSGATIR 200  
 241 QRDFTAFPRMLTTRGHGPHDGSCLSPVRLQISTLANELRDSRLANDRDSGATIR 300  
 201 IGIYAGICAGLALIFGALIFKWSHSEKIQNLISLANLPSSGLANAAVEGIRS 260  
 301 IGIYAGICAGLALIFGALIFKWSHSEKIQNLISLANLPSSGLANAAVEGIRS 360  
 261 EENITYTEENYVEVEENYCYVSSROQSPQRCPEFAP 301  
 361 EENITYTEENYVEVEENYCYVSSROQSPQRCPEFAP 401

Search completed: November 22, 2003, 05:40:21  
 Job time : 59.9588 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 05:38:42 ; Search time 19.6529 Seconds  
(without alignments)  
648.024 Million cell updates/sec

Title: US-10-004-633-24

Perfect score: 1587

Sequence: 1 MFSHLPDVCVLLLLLLTR.....CYVSSROOPSQPLCRFPAMP 301

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents\_AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	1587	100.0	301	US-08-829-525-24	Sequence 24, Appli
2	1587	100.0	301	US-08-609-583A-24	Sequence 24, Appli
3	1587	100.0	301	US-08-937-359-24	Sequence 24, Appli
4	1587	100.0	301	US-09-310-367-24	Sequence 24, Appli
5	1587	100.0	301	US-09-033-337-24	Sequence 24, Appli
6	1587	100.0	301	US-09-464-231-24	Sequence 24, Appli
7	883	55.6	281	US-08-487-748A-9	Sequence 9, Appli
8	883	55.6	281	US-08-487-748A-10	Sequence 10, Appli
9	883	55.6	281	US-08-480-070C-10	Sequence 10, Appli
10	883	55.6	281	US-08-829-525-10	Sequence 10, Appli
11	883	55.6	281	US-08-609-583A-10	Sequence 10, Appli
12	883	55.6	281	US-08-937-359-10	Sequence 10, Appli
13	883	55.6	281	US-09-310-367-10	Sequence 10, Appli
14	883	55.6	281	US-09-033-337-10	Sequence 10, Appli
15	883	55.6	281	US-09-464-231-10	Sequence 10, Appli
16	256	16.1	451	US-08-287-001A-2	Sequence 2, Appli
17	256	16.1	451	PCT-US95-09941-2	Sequence 2, Appli
18	133.5	8.4	365	US-08-928-383B-23	Sequence 23, Appli
19	133.5	8.4	365	US-08-928-383B-24	Sequence 24, Appli
20	133.5	8.3	387	US-09-175-928-2	Sequence 2, Appli
21	131.5	8.3	365	US-08-928-383B-2	Sequence 2, Appli
22	131.5	8.3	365	US-08-928-383B-26	Sequence 26, Appli
23	131.5	8.3	581	US-08-724-394A-3	Sequence 3, Appli
24	126	7.9	347	US-09-667-135-4	Sequence 4, Appli
25	124.5	7.8	352	US-09-996-243-505	Sequence 505, App
26	124.5	7.8	365	US-08-979-424-3	Sequence 3, Appli
27	124.5	7.8	365	US-09-272-496-2	Sequence 2, Appli

28	120	7.6	581	2	US-08-724-394A-2	Sequence 2, Appli
29	118.5	7.5	398	3	US-09-189-035-6	Sequence 6, Appli
30	118.5	7.5	398	3	US-09-382-086-6	Sequence 6, Appli
31	118.5	7.5	398	4	US-08-999-689A-5	Sequence 5, Appli
32	118	7.4	322	4	US-09-667-135-2	Sequence 2, Appli
33	115.5	7.3	466	4	US-09-604-107A-8	Sequence 8, Appli
34	112.5	7.1	319	1	US-08-597-495B-22	Sequence 22, Appli
35	112.5	7.1	319	3	US-09-068-051A-22	Sequence 22, Appli
36	112.5	7.1	319	4	US-09-336-536-67	Sequence 67, Appli
37	112.5	7.1	319	4	US-09-254-465A-6	Sequence 6, Appli
38	110	6.9	373	4	US-09-926-243-503	Sequence 503, App
39	108	6.8	270	4	US-09-254-465A-24	Sequence 24, Appli
40	108	6.8	273	4	US-09-254-465A-26	Sequence 26, Appli
41	108	6.8	489	4	US-09-667-135-10	Sequence 8, Appli
42	108	6.8	1138	1	US-08-323-474-8	Sequence 8, Appli
43	108	6.8	1138	2	US-08-469-537A-98	Sequence 98, Appli
44	108	6.8	1138	2	US-08-220-240A-5	Sequence 5, Appli
45	105.5	6.6	318	3	US-09-068-051A-32	Sequence 32, Appli

## ALIGNMENTS

RESULT 1.  
US-08-829-525-24  
Sequence 24, Application US/08829525

Patent No. 6084083

GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036/2721

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/829,525

FILING DATE: 28-MAR-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/609,583

FILING DATE: 01-MAR-1996

APPLICATION NUMBER: US 08/487,748

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/398,633

FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7853-081

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 301 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

US-08-829-525-24

Query Match 100.0%; Score 1587; DB 3; Length 301;  
Best Local Similarity 100.0%; Pred. No. 9.3e-155;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESHLPFDCVLLLLLLLLLRSSSEVEYRAEVGNAYLPCFYTPAAPGNLVPVCMGKACPV 60  
1 MSHLPFDCVLLLLLLLLLRSSSEVEYRAEVGNAYLPCFYTPAAPGNLVPVCMGKACPV 60  
DB 1 MSHLPFDCVLLLLLLLLLRSSSEVEYRAEVGNAYLPCFYTPAAPGNLVPVCMGKACPV 60  
QY 61 FEGGNVLTDERDVNWTSRWLNDFKRGDVSLLTEENVTLADSGIYCCRIQIPIGIMND 120  
61 FEGGNVLTDERDVNWTSRWLNDFKRGDVSLLTEENVTLADSGIYCCRIQIPIGIMND 120  
DB 122 EKFNKLTVIKPAKVTAPTLQRFDAFPFRLTTTGHGPAETQTLGSLPDINLTQISTLA 180  
121 EKFNKLTVIKPAKVTAPTLQRFDAFPFRLTTTGHGPAETQTLGSLPDINLTQISTLA 180  
QY 181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALLFKWYSHSKEKIONLSLI 240  
181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALLFKWYSHSKEKIONLSLI 240  
DB 241 SLANLPSPGLANAVAGIRSEENIYTIENNVYVEEPEPNYYCYVSSROQPSQPLGCRFAM 300  
241 SLANLPSPGLANAVAGIRSEENIYTIENNVYVEEPEPNYYCYVSSROQPSQPLGCRFAM 300  
QY 301 P 301  
301 P 301  
DB 301 P 301

## RESULT 2

US-08-609-583A-24  
; Sequence 24, Application US/08609583A  
; Patent No. 6204371

; GENERAL INFORMATION:

; APPLICANT: Levinson, Douglas A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036/2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US 08/609,583A

; FILING DATE: 01-MAR-1996

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/487,748

; FILING DATE: 07-JUN-1995

; APPLICATION NUMBER: US 08/398,633

; FILING DATE: 03-MAR-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7853-048

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090

; TELEFAX: 212-869-8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 301 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: Protein  
; FRAGMENT TYPE: Internal  
US-08-609-583A-24

Query Match 100.0%; Score 1587; DB 3; Length 301;  
Best Local Similarity 100.0%; Pred. No. 9.3e-155;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESHLPFDCVLLLLLLLLLRSSSEVEYRAEVGNAYLPCFYTPAAPGNLVPVCMGKACPV 60  
1 MSHLPFDCVLLLLLLLLLRSSSEVEYRAEVGNAYLPCFYTPAAPGNLVPVCMGKACPV 60  
DB 1 MSHLPFDCVLLLLLLLLLRSSSEVEYRAEVGNAYLPCFYTPAAPGNLVPVCMGKACPV 60  
QY 61 FEGGNVLTDERDVNWTSRWLNDFKRGDVSLLTEENVTLADSGIYCCRIQIPIGIMND 120  
61 FEGGNVLTDERDVNWTSRWLNDFKRGDVSLLTEENVTLADSGIYCCRIQIPIGIMND 120  
DB 61 FEGGNVLTDERDVNWTSRWLNDFKRGDVSLLTEENVTLADSGIYCCRIQIPIGIMND 120  
121 EKFNKLTVIKPAKVTAPTLQRFDAFPFRLTTTGHGPAETQTLGSLPDINLTQISTLA 180  
121 EKFNKLTVIKPAKVTAPTLQRFDAFPFRLTTTGHGPAETQTLGSLPDINLTQISTLA 180  
QY 181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALLFKWYSHSKEKIONLSLI 240  
181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALLFKWYSHSKEKIONLSLI 240  
DB 241 SLANLPSPGLANAVAGIRSEENIYTIENNVYVEEPEPNYYCYVSSROQPSQPLGCRFAM 300  
241 SLANLPSPGLANAVAGIRSEENIYTIENNVYVEEPEPNYYCYVSSROQPSQPLGCRFAM 300  
QY 301 P 301  
301 P 301  
DB 301 P 301

## RESULT 3

US-08-937-399-24

; Sequence 24, Application US/08937399

; Patent No. 6288218

; GENERAL INFORMATION:

; APPLICANT: Levinson, Douglas A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036/2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US 08/937,399

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/609,583

; FILING DATE: 01-MAR-1996

; APPLICATION NUMBER: US 08/487,748

; FILING DATE: 07-JUN-1995

; APPLICATION NUMBER: US 08/398,633

; FILING DATE: 03-MAR-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7853-048

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090

; TELEFAX: 212-869-8864

; TELEX: 66141 PENNIE

;; INFORMATION FOR SEQ ID NO: 24:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 301 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: internal  
;; US-08-937-399-24

Query Match 100.0%; Score 1587; DB 3; Length 301;  
Best Local Similarity 100.0%; Pred. No. 9.3e-155;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSLPDCVLLLLLLLLTSSSEVEYRAVGONAYLPCFTTAPAPGNLVPVCMKGACPV 60  
DB 1 MFSLPDCVLLLLLLLLTSSSEVEYRAVGONAYLPCFTTAPAPGNLVPVCMKGACPV 60

QY 61 PEGCNVVLRTDERDVNWTSTRYMLNGDFRKGDVSLTIENVTLADSGIYCCHIQIPGIMND 120  
DB 61 PEGCNVVLRTDERDVNWTSTRYMLNGDFRKGDVSLTIENVTLADSGIYCCHIQIPGIMND 120

QY 121 EKFNKLVIRPAKVTPTLQDFTTAPPRMLTTRGHGPAETQTLGSLPDINLTQISTLA 180  
DB 121 EKFNKLVIRPAKVTPTLQDFTTAPPRMLTTRGHGPAETQTLGSLPDINLTQISTLA 180

QY 181 NELDSRLANDLRDSGATIRIGIYIGAGICAGLALIFGALIFKWSHSEKIONLSLI 240  
DB 181 NELDSRLANDLRDSGATIRIGIYIGAGICAGLALIFGALIFKWSHSEKIONLSLI 240

QY 241 SLANLPPSGLANAAVEGIRSEENITYTIEENVYEEBEPNEYCYVSSROQPSQPLGCRFAM 300  
DB 241 SLANLPPSGLANAAVEGIRSEENITYTIEENVYEEBEPNEYCYVSSROQPSQPLGCRFAM 300

QY 301 P 301  
DB 301 P 301

RESULT 4  
US-09-310-367-24  
; Sequence 24, Application US/09310367  
; Patent No. 6414117  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/310.367  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/829.525  
; FILING DATE: 28-MAR-1997  
; APPLICATION NUMBER: US 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Coruzzi, Laura A.  
;; REGISTRATION NUMBER: 30,742  
;; REFERENCE/DOCKET NUMBER: 7853-081  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-790-9090  
;; TELEFAX: 212-869-8864  
;; TELE: 66141 PENNIE  
;; INFORMATION FOR SEQ ID NO: 24:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 301 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: internal  
;; US-09-310-367-24

Query Match 100.0%; Score 1587; DB 4; Length 301;  
Best Local Similarity 100.0%; Pred. No. 9.3e-155;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSLPDCVLLLLLLLLTSSSEVEYRAVGONAYLPCFTTAPAPGNLVPVCMKGACPV 60  
DB 1 MFSLPDCVLLLLLLLLTSSSEVEYRAVGONAYLPCFTTAPAPGNLVPVCMKGACPV 60

QY 61 PEGCNVVLRTDERDVNWTSTRYMLNGDFRKGDVSLTIENVTLADSGIYCCHIQIPGIMND 120  
DB 61 PEGCNVVLRTDERDVNWTSTRYMLNGDFRKGDVSLTIENVTLADSGIYCCHIQIPGIMND 120

QY 121 EKFNKLVIRPAKVTPTLQDFTTAPPRMLTTRGHGPAETQTLGSLPDINLTQISTLA 180  
DB 121 EKFNKLVIRPAKVTPTLQDFTTAPPRMLTTRGHGPAETQTLGSLPDINLTQISTLA 180

QY 181 NELDSRLANDLRDSGATIRIGIYIGAGICAGLALIFGALIFKWSHSEKIONLSLI 240  
DB 181 NELDSRLANDLRDSGATIRIGIYIGAGICAGLALIFGALIFKWSHSEKIONLSLI 240

QY 241 SLANLPPSGLANAAVEGIRSEENITYTIEENVYEEBEPNEYCYVSSROQPSQPLGCRFAM 300  
DB 241 SLANLPPSGLANAAVEGIRSEENITYTIEENVYEEBEPNEYCYVSSROQPSQPLGCRFAM 300

QY 301 P 301  
DB 301 P 301

RESULT 5  
US-09-032-337-24  
; Sequence 24, Application US/09032337  
; Patent No. 6455685  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/032.337  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-032-337-24

Query Match 100.0%; Score 1587; DB 4; Length 301;  
Best local Similarity 100.0%; Pred. No. 9.3e-155;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESHLPDVCVLLLLLLLLLRSSEVEYRAEVGQNAVLPCEYTPAPAGNLVPCWGKACPV 60  
DB 1 MESHLPDVCVLLLLLLLLLRSSEVEYRAEVGQNAVLPCEYTPAPAGNLVPCWGKACPV 60  
QY 61 FECCGVNLTDRDENVYTSRYWLNQDFRKGVSLENTVLTADSGIYCCRIQIPIGIMND 120  
DB 61 FECCGVNLTDRDENVYTSRYWLNQDFRKGVSLENTVLTADSGIYCCRIQIPIGIMND 120  
QY 121 EKENLKVIPKAPYTPAPLQDFTAAPRMLTTTGHGPAETOTLSLPDINTQISTLA 180  
DB 121 EKENLKVIPKAPYTPAPLQDFTAAPRMLTTTGHGPAETOTLSLPDINTQISTLA 180  
QY 181 NEIRDSRLANDLRDSGATTIRIGIYGAGICAGLALALIFGALIFKWSHSEKIQNLSTLI 240  
DB 181 NEIRDSRLANDLRDSGATTIRIGIYGAGICAGLALALIFGALIFKWSHSEKIQNLSTLI 240  
QY 241 SLANLPPSGLANVAAGIRSEENITYTEENVYEEPEENYYCVSSROOPSQPLGCRFAM 300  
DB 241 SLANLPPSGLANVAAGIRSEENITYTEENVYEEPEENYYCVSSROOPSQPLGCRFAM 300  
QY 301 P 301  
DB 301 P 301

RESULT 6  
US-09-464-231-24  
Sequence 24, Application US/09464231  
Patent No. 6562343  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/464,231  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-09-464-231-24

Query Match 100.0%; Score 1587; DB 4; Length 301;  
Best local Similarity 100.0%; Pred. No. 9.3e-155;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESHLPDVCVLLLLLLLLLRSSEVEYRAEVGQNAVLPCEYTPAPAGNLVPCWGKACPV 60  
DB 1 MESHLPDVCVLLLLLLLLLRSSEVEYRAEVGQNAVLPCEYTPAPAGNLVPCWGKACPV 60  
QY 61 FECCGVNLTDRDENVYTSRYWLNQDFRKGVSLENTVLTADSGIYCCRIQIPIGIMND 120  
DB 61 FECCGVNLTDRDENVYTSRYWLNQDFRKGVSLENTVLTADSGIYCCRIQIPIGIMND 120  
QY 121 EKENLKVIPKAPYTPAPLQDFTAAPRMLTTTGHGPAETOTLSLPDINTQISTLA 180  
DB 121 EKENLKVIPKAPYTPAPLQDFTAAPRMLTTTGHGPAETOTLSLPDINTQISTLA 180  
QY 181 NEIRDSRLANDLRDSGATTIRIGIYGAGICAGLALALIFGALIFKWSHSEKIQNLSTLI 240  
DB 181 NEIRDSRLANDLRDSGATTIRIGIYGAGICAGLALALIFGALIFKWSHSEKIQNLSTLI 240  
QY 241 SLANLPPSGLANVAAGIRSEENITYTEENVYEEPEENYYCVSSROOPSQPLGCRFAM 300  
DB 241 SLANLPPSGLANVAAGIRSEENITYTEENVYEEPEENYYCVSSROOPSQPLGCRFAM 300  
QY 301 P 301  
DB 301 P 301

RESULT 7  
US-08-487-748A-9  
Sequence 9, Application US/08487748A  
Patent No. 5721351  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.

ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,748A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-487-748A-9

Query Match 55.6%; Score 883; DB 1; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFDVLLLLLLLTSSSEVEYRAVGNAYLPCEYTPAAGNLVPCWKGACPV 60  
DB 1 MFGSLTNCVLLLLQLLARSLEDGYKVEGKNAVLPESYTLPTSGTLVPMCKGFCPW 60  
QY 61 FEGGNVLRTERDVNY-TSRYWNGDFRKGDVSLTENTVTLADSGIYCCRIQPIGMN 119  
DB 61 SQCTNELLRTERDVNYKSSRYQLKGLNKGDVSLIKNTLDDHGYCCRIQPIGMN 120  
QY 120 DEKNLKIIVAKVTPAPTLQDFTAPFPMLTTRGHGPAETOTIGSLPDLNLTQISTL 179  
DB 121 DKLELKLDIAKAVTPAQTHAGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELRDSRLANDSGATIRIGIYGAGICAGLALALIFGALIFKWTSHSEKIQNLSTL 239  
DB 180 ADE-----IKDSGETIRTAIHIGVGSAGLTLALIIIGVLIKWISCKKKLSISL 230  
QY 240 ISLANLPSSGLANAVABGIRSEENIYTIENVEVEEENEYCYVSSRQOPS 291  
DB 231 ITLANLPSSGLANAGAVRIRSEENIYTIENVEVEEENEYCYVNS-QOPS 281

RESULT 8  
US-08-487-748A-10  
Sequence 10, Application US/08487748A  
Patent No. 5721351  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,748A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-487-748A-10

Query Match 55.6%; Score 883; DB 1; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFDVLLLLLLLTSSSEVEYRAVGNAYLPCEYTPAAGNLVPCWKGACPV 60  
DB 1 MFGSLTNCVLLLLQLLARSLEDGYKVEGKNAVLPESYTLPTSGTLVPMCKGFCPW 60  
QY 61 FEGGNVLRTERDVNY-TSRYWNGDFRKGDVSLTENTVTLADSGIYCCRIQPIGMN 119  
DB 61 SQCTNELLRTERDVNYKSSRYQLKGLNKGDVSLIKNTLDDHGYCCRIQPIGMN 120  
QY 120 DEKNLKIIVAKVTPAPTLQDFTAPFPMLTTRGHGPAETOTIGSLPDLNLTQISTL 179  
DB 121 DKLELKLDIAKAVTPAQTHAGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELRDSRLANDSGATIRIGIYGAGICAGLALALIFGALIFKWTSHSEKIQNLSTL 239  
DB 180 ADE-----IKDSGETIRTAIHIGVGSAGLTLALIIIGVLIKWISCKKKLSISL 230  
QY 240 ISLANLPSSGLANAVABGIRSEENIYTIENVEVEEENEYCYVSSRQOPS 291  
DB 231 ITLANLPSSGLANAGAVRIRSEENIYTIENVEVEEENEYCYVNS-QOPS 281

RESULT 9  
US-08-480-070C-10  
Sequence 10, Application US/08480070C  
Patent No. 6066498  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,070C  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-024

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-480-070C-10

Query Match 55.6%; Score 883; DB 3; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;

Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFDVLLLLLLLTSSSEVYRAVGNAYLPCEYTPAPGNLVPVCMGKACPV 60  
DB 1 MFGSLTNCVLLLLQLLARSLEDDYKVEVGNAYLPSCYTLPTSGLVPMCMGKFCPW 60  
QY 61 FEGGNVLTDERDVNYM-TSRVWNGDFPKGVSLLTENTLADSGIYCCRIQIPGIN 119  
DB 61 SGTNELLTDERNVYOKSSRYOLKGDINKGVSLIKNVTLDDHGYCCRIQIPGLMN 120  
QY 120 DEKENLKVYKPAKTPAPTLQDFTPAAPRMLTTRGHGPAETQTLGSLPDINTQISTL 179  
DB 121 DKLEIKLDIKAAKTPACTAGDSTTASPRITLTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELRDSRLANDLRDSCATIRIGIYGAGICAGLALALIFGALIFKWSHSEKIKQNTSL 239  
DB 180 ADE-----IKDGETIRTAIHIGVGSAGLTALIGVLLKWSCKKKLSLSL 230  
QY 240 ISLANIPPSGLANAVABGIRSEENITYTEENVYEEBPNYYCYVSSROOPS 291  
DB 231 ITLANIPPSGLANAGAVRIRSEENITYTEENVYEEBPNYYCYVNS-CQPS 281

## RESULT 10

US-08-829-525-10  
Sequence 10, Application US/08829525  
Patent No. 6084083  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/829,525  
FILING DATE: 28-MAR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-081

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-829-525-10

Query Match 55.6%; Score 883; DB 3; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;

Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFDVLLLLLLLTSSSEVYRAVGNAYLPCEYTPAPGNLVPVCMGKACPV 60  
DB 1 MFGSLTNCVLLLLQLLARSLEDDYKVEVGNAYLPSCYTLPTSGLVPMCMGKFCPW 60  
QY 61 FEGGNVLTDERDVNYM-TSRVWNGDFPKGVSLLTENTLADSGIYCCRIQIPGIN 119  
DB 61 SGTNELLTDERNVYOKSSRYOLKGDINKGVSLIKNVTLDDHGYCCRIQIPGLMN 120  
QY 120 DEKENLKVYKPAKTPAPTLQDFTPAAPRMLTTRGHGPAETQTLGSLPDINTQISTL 179  
DB 121 DKLEIKLDIKAAKTPACTAGDSTTASPRITLTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELRDSRLANDLRDSCATIRIGIYGAGICAGLALALIFGALIFKWSHSEKIKQNTSL 239  
DB 180 ADE-----IKDGETIRTAIHIGVGSAGLTALIGVLLKWSCKKKLSLSL 230  
QY 240 ISLANIPPSGLANAVABGIRSEENITYTEENVYEEBPNYYCYVSSROOPS 291  
DB 231 ITLANIPPSGLANAGAVRIRSEENITYTEENVYEEBPNYYCYVNS-CQPS 281

## RESULT 11

US-08-609-583A-10  
Sequence 10, Application US/08609583A  
Patent No. 6204371  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583A  
FILING DATE: 01-MAR-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090



TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-609-583A-10

Query Match 55.6%; Score 883; DB 3; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFCVLLLLLLLTSSSEVEYRAEYVGNAYLPFCYTPAAGNLPVPCWKGACPV 60  
DB 1 MFSGLTLCVLLLLQLLARSLEDGYKVEGKNAVLPSCYTLPTSGTLVPKMGWGFPCW 60  
QY 61 FECCGVVLRTERDVNYW-TSRVWLNCDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMN 119  
DB 61 SQCTNELRTDERNVYTKSSRYQLKGDVSLTIKNTLDDHGYCCRIQIPGIMN 120  
QY 120 DEKNLKVIRPAKVTAPTLQDFTAAFPKMLTTRGHGPAETQTLGSLPDINTQISTL 179  
DB 121 DKKELKLDIKAKVTAPQTAHGDSTASPTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELRDSRLANDLRSGATIRIGIYGAGICAGLALIFGALLFKWYSHSEKIQNSTL 239  
DB 180 ADE-----IKDSGETIRTAIHIGVSAAGLTALITGLILKWSCKKKKLSLSL 230  
QY 240 ISLANLPPSGLANAABEGIRSEENIYTIENVEVEEPEENEYCYVSSRQPS 291  
DB 231 ITLANLPPGGLANAGAVIRSEENIYTIENVEVEENENYCYVNS-QQPS 281

## RESULT 12

US-08-937-399-10  
Sequence 10, Application US/08937399  
Patent No. 6288218

## GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/937,399  
FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-937-399-10

Query Match 55.6%; Score 883; DB 3; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFCVLLLLLLLTSSSEVEYRAEYVGNAYLPFCYTPAAGNLPVPCWKGACPV 60  
DB 1 MFSGLTLCVLLLLQLLARSLEDGYKVEGKNAVLPSCYTLPTSGTLVPKMGWGFPCW 60  
QY 61 FECCGVVLRTERDVNYW-TSRVWLNCDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMN 119  
DB 61 SQCTNELRTDERNVYTKSSRYQLKGDVSLTIKNTLDDHGYCCRIQIPGIMN 120  
QY 120 DEKNLKVIRPAKVTAPTLQDFTAAFPKMLTTRGHGPAETQTLGSLPDINTQISTL 179  
DB 121 DKKELKLDIKAKVTAPQTAHGDSTASPTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELRDSRLANDLRSGATIRIGIYGAGICAGLALIFGALLFKWYSHSEKIQNSTL 239  
DB 180 ADE-----IKDSGETIRTAIHIGVSAAGLTALITGLILKWSCKKKKLSLSL 230  
QY 240 ISLANLPPSGLANAABEGIRSEENIYTIENVEVEEPEENEYCYVSSRQPS 291  
DB 231 ITLANLPPGGLANAGAVIRSEENIYTIENVEVEENENYCYVNS-QQPS 281

## RESULT 13

US-09-310-367-10  
Sequence 10, Application US/09310367  
Patent No. 6414117

## GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/310,367  
FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/829,525  
FILING DATE: 28-MAR-1997  
APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-081

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-310-367-10

Query Match 55.6%; Score 883; DB 4; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MESHLPDVCVLLLLLLLTSSSEVEYRAEYVGNAYLPCCYTPAAPGNLVPCMGKACPV 60  
DB 1 MMSGTLNLCVLLLLQLLARSLEDGYKVEGKNAVLPSCYTLPTSGLTVMCMGKGFPCW 60  
QY 61 PFCGNVLTDERDVNYW-TSRVWLNDFPKGVSILTENTVLADSGIYCCRIQIPGIN 119  
DB 61 SQTNELRTDERNVYQKSSRYQLKGDINKGVSLIKNVTLDDHGTCCRIQFPGLMN 120  
QY 120 DEKFNKLVIKPAKVPAPTLQDFTAPFPRMLTTRGHGPAETQTLGSLPDINLTQISTL 179  
DB 121 DKLELKDIDKAKVPAQTAGDSTTASPTLTTERNG-SETQTLVTLHNNNGTISTW 179  
QY 180 ANELRSRLANDLRDGGATIRIGIYGAGICAGLALALIFGALIFKWSHSEKIONLSL 239  
DB 180 ADE-----IKDSETRIRTAIHIGVVSAGLTALILIGVILKWSCKKKLSLSL 230  
QY 240 ISLANLPSCGLANAVABGIRSEENITYTTEENYVEEENPYCYVSSROOPS 291  
DB 231 ITLANLPFGLANAGAVRIRSEENITYTTEENYVEEENPYCYVNS-CQPS 281

RESULT 14  
US-09-032-337-10  
Sequence 10, Application US/09032337  
Patent No. 6455685  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Penite & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/032,337  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-016

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-032-337-10

Query Match 55.6%; Score 883; DB 4; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MESHLPDVCVLLLLLLLTSSSEVEYRAEYVGNAYLPCCYTPAAPGNLVPCMGKACPV 60  
DB 1 MMSGTLNLCVLLLLQLLARSLEDGYKVEGKNAVLPSCYTLPTSGLTVMCMGKGFPCW 60  
QY 61 PFCGNVLTDERDVNYW-TSRVWLNDFPKGVSILTENTVLADSGIYCCRIQIPGIN 119  
DB 61 SQTNELRTDERNVYQKSSRYQLKGDINKGVSLIKNVTLDDHGTCCRIQFPGLMN 120  
QY 120 DEKFNKLVIKPAKVPAPTLQDFTAPFPRMLTTRGHGPAETQTLGSLPDINLTQISTL 179  
DB 121 DKLELKDIDKAKVPAQTAGDSTTASPTLTTERNG-SETQTLVTLHNNNGTISTW 179  
QY 180 ANELRSRLANDLRDGGATIRIGIYGAGICAGLALALIFGALIFKWSHSEKIONLSL 239  
DB 180 ADE-----IKDSETRIRTAIHIGVVSAGLTALILIGVILKWSCKKKLSLSL 230  
QY 240 ISLANLPSCGLANAVABGIRSEENITYTTEENYVEEENPYCYVSSROOPS 291  
DB 231 ITLANLPFGLANAGAVRIRSEENITYTTEENYVEEENPYCYVNS-CQPS 281

RESULT 15  
US-09-464-231-10  
Sequence 10, Application US/09464231  
Patent No. 6562343  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Penite & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/464,231  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-790-9090  
 TELEFAX: 212-869-8864  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 281 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 US-09-464-231-10

Query Match 55.6%; Score 883; DB 4; Length 281;  
 Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
 Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MESHLPEDCVLLLLLLLTSSSEVEYRAVGONATLPCTTPAPGNLVPVCGKACPV 60  
 Db 1 MESSGLTINCIVLLLOLLARSLDEGYKVEVGKNAYLPCSYTLPTSGTLVPMCGKGCPCW 60  
 QY 61 FECSNVVLRTPEDVDYVW-TSRVWNGDFPRKGDVSLTENTLADSGIYCCRIQIPIGN 119  
 Db 61 SQTNEELRTDBRNVTYKSSRYQKGDVSLIKNVTLDHGTCCRIQPPGLMN 120  
 QY 120 DEKFLKLVIRPAKVPAPTLQDPTAAPPMLTTRGHGPAETQTLSLPDINLTQISTL 179  
 Db 121 DKKELKLDIKAAKVTPAQTAAHGDSTTASPRITLTERNG-SETQTLVTLHNNNGTKISTW 179  
 QY 180 ANELRDSRLANDLRDSGATIRIGIYGAGICAGLALIFGALIFKWTSHSKEKIQNLSTL 239  
 Db 180 ADE-----IKDSGETIRTAIHIGVGSAGLTALIGVLTKWYSCKKKKLSISL 230  
 QY 240 ISLANLPPSGLANAVAEGRSESENYTTEENYVEEENPYCYVSSROOPS 291  
 Db 231 ITLANLPPGGLANAGAVRIRSESENYTTEENYVEEENPYCYVNS-QQPS 281

Search completed: November 22, 2003, 05:45:21  
 Job time : 21.6529 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 22, 2003, 05:43:31 ; Search time 44.9948 Seconds  
(without alignments)  
1221.260 Million cell updates/sec

Title: US-10-004-633-24

Perfect score: 1587

Sequence: 1 MFSLPFDVCLLLLLLTTR.....CYVSRQRPQLGCRFAMP 301

## Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

## Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1587	100.0	301	US-10-252-131-24	Sequence 24, Appl
2	1587	100.0	301	US-10-004-633-24	Sequence 24, Appl
3	1587	100.0	301	US-10-188-012-31	Sequence 31, Appl
4	1581	99.6	301	US-10-188-012-29	Sequence 29, Appl
5	940	59.2	185	US-09-528-063-34	Sequence 34, Appl
6	940	59.2	185	US-10-414-378-34	Sequence 34, Appl
7	885	55.8	281	US-10-188-012-11	Sequence 11, Appl
8	883	55.6	281	US-10-252-131-10	Sequence 10, Appl
9	883	55.6	281	US-10-004-633-10	Sequence 9, Appl
10	883	55.6	281	US-10-188-012-9	Sequence 9, Appl
11	718	45.2	142	US-09-966-546-22	Sequence 22, Appl
12	718	45.2	142	US-09-966-545-22	Sequence 22, Appl
13	718	45.2	142	US-09-965-212-22	Sequence 22, Appl
14	718	45.2	142	US-10-189-940-22	Sequence 22, Appl
15	718	45.2	142	US-10-189-940-146	Sequence 146, Appl

16	613	38.6	125	16	US-10-189-940-145	Sequence 145, App
17	333	21.0	282	15	US-10-188-012-3	Sequence 3, Appl
18	323.5	20.4	305	15	US-10-188-012-1	Sequence 1, Appl
19	322	20.3	359	15	US-10-188-012-17	Sequence 17, Appl
20	322	20.3	359	15	US-10-188-012-19	Sequence 19, Appl
21	322	20.3	359	15	US-10-188-012-23	Sequence 23, Appl
22	319.5	20.1	364	15	US-10-188-012-25	Sequence 25, Appl
23	319	20.1	365	15	US-10-188-012-21	Sequence 21, Appl
24	317.5	20.0	364	15	US-10-188-012-27	Sequence 27, Appl
25	273.5	17.2	305	15	US-10-188-012-5	Sequence 5, Appl
26	273.5	17.2	305	15	US-10-188-012-7	Sequence 7, Appl
27	250.5	15.8	345	15	US-10-188-012-13	Sequence 13, Appl
28	246.5	15.5	345	15	US-10-188-012-15	Sequence 15, Appl
29	245.5	15.5	378	15	US-10-188-012-35	Sequence 35, Appl
30	245.5	15.5	378	15	US-10-188-012-35	Sequence 35, Appl
31	245.5	15.5	379	11	US-09-813-153-138	Sequence 138, App
32	238.5	15.0	183	9	US-09-739-907-65	Sequence 65, App
33	132.5	8.3	387	14	US-10-114-893-133	Sequence 133, App
34	132.5	8.3	387	15	US-10-016-249-2	Sequence 2, Appl
35	131.5	8.3	523	10	US-09-910-174A-11	Sequence 11, Appl
36	131.5	8.3	523	10	US-09-955-866-10	Sequence 10, Appl
37	131.5	8.3	523	10	US-09-896-738-16	Sequence 16, Appl
38	131.5	8.3	523	12	US-10-087-887-53	Sequence 53, Appl
39	127.5	8.0	365	9	US-09-899-634A-4	Sequence 4, Appl
40	126	7.9	347	15	US-10-281-478-9	Sequence 9, Appl
41	125	7.9	505	12	US-10-114-153-12	Sequence 12, Appl
42	124.5	7.8	261	9	US-09-889-634A-2	Sequence 2, Appl
43	124.5	7.8	352	9	US-09-989-722-505	Sequence 505, App
44	124.5	7.8	352	9	US-09-989-723-505	Sequence 505, App
45	124.5	7.8	352	9	US-09-989-279-505	Sequence 505, App

## ALIGNMENTS

RESULT 1  
US-10-252-131-24

Sequence 24, Application US/10252131  
Publication No US20030158399A1

GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSER: Pennie &amp; Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036/2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/252,131

FILING DATE: 20-Sep-2002

CLASSIFICATION: &lt;unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/609,583

FILING DATE: 01-MAR-1996

APPLICATION NUMBER: US 08/487,748

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/398,633

FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Cornuzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7853-048

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 24:  
US-10-252-131-24

Query Match 100.0%; Score 1587; DB 12; Length 301;  
Best Local Similarity 100.0%; Pred. No. 2.7e-151;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSLPFDVLLLLLLLTSSEVEYRAEYVGNAYLPCFYTPAAGNLVPCWGKACPV 60  
DB 1 MFSLPFDVLLLLLLLTSSEVEYRAEYVGNAYLPCFYTPAAGNLVPCWGKACPV 60  
QY 61 FECGNVVLRTDERDVNWTSRWYLNQDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
DB 61 FECGNVVLRTDERDVNWTSRWYLNQDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
QY 121 EKFNKLVIKPAKVTAPTLQRFPAAPRMLTTRGHGPAETQTLSLPDINTQISTLA 180  
DB 121 EKFNKLVIKPAKVTAPTLQRFPAAPRMLTTRGHGPAETQTLSLPDINTQISTLA 180  
QY 181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNLSLI 240  
DB 181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNLSLI 240  
QY 241 SLANLPPSGLANA VAEGIRSEENIYTIENNVYVEEPEENYYCYVSSROQPSQPLGCRFAM 300  
DB 241 SLANLPPSGLANA VAEGIRSEENIYTIENNVYVEEPEENYYCYVSSROQPSQPLGCRFAM 300  
QY 301 P 301  
DB 301 P 301

RESULT 2  
US-10-004-633-24

/ Sequence 24, Application US/10004633  
/ Publication No. US20030069196A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Levinson, Douglas A.  
/ APPLICANT: Lloyd, Clare M.  
/ APPLICANT: McCarthy, Sean A.  
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
/ FILE REFERENCE: 7853-125  
/ CURRENT APPLICATION NUMBER: US/10/004,633  
/ CURRENT FILING DATE: 2001-12-04  
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986  
/ PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583  
/ PRIOR FILING DATE: EARLIER FILING DATE: 1996-03-01  
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/487,748  
/ PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07  
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/398,633  
/ PRIOR FILING DATE: EARLIER FILING DATE: 1995-03-05  
/ NUMBER OF SEQ ID NOS: 49  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 24  
/ LENGTH: 301  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
US-10-004-633-24

Query Match 100.0%; Score 1587; DB 15; Length 301;  
Best Local Similarity 100.0%; Pred. No. 2.7e-151;

Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSLPFDVLLLLLLLTSSEVEYRAEYVGNAYLPCFYTPAAGNLVPCWGKACPV 60  
DB 1 MFSLPFDVLLLLLLLTSSEVEYRAEYVGNAYLPCFYTPAAGNLVPCWGKACPV 60  
QY 61 FECGNVVLRTDERDVNWTSRWYLNQDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
DB 61 FECGNVVLRTDERDVNWTSRWYLNQDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
QY 121 EKFNKLVIKPAKVTAPTLQRFPAAPRMLTTRGHGPAETQTLSLPDINTQISTLA 180  
DB 121 EKFNKLVIKPAKVTAPTLQRFPAAPRMLTTRGHGPAETQTLSLPDINTQISTLA 180  
QY 181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNLSLI 240  
DB 181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNLSLI 240  
QY 241 SLANLPPSGLANA VAEGIRSEENIYTIENNVYVEEPEENYYCYVSSROQPSQPLGCRFAM 300  
DB 241 SLANLPPSGLANA VAEGIRSEENIYTIENNVYVEEPEENYYCYVSSROQPSQPLGCRFAM 300  
QY 301 P 301  
DB 301 P 301

RESULT 3  
US-10-188-012-31

/ Sequence 31, Application US/10188012  
/ Publication No. US20030124114A1  
/ GENERAL INFORMATION:  
/ APPLICANT: McIntire, Jennifer Jones  
/ APPLICANT: Umetsu, Dale T.  
/ APPLICANT: Dekruyf, Rosemarie  
/ APPLICANT: Kuchroo, Vijay  
/ APPLICANT: Freeman, Gordon J.  
/ TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
/ FILE REFERENCE: STAN-235  
/ CURRENT APPLICATION NUMBER: US/10/188,012  
/ CURRENT FILING DATE: 2002-07-01  
/ PRIOR APPLICATION NUMBER: 60/302,344  
/ PRIOR FILING DATE: 2001-06-29  
/ NUMBER OF SEQ ID NOS: 36  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 31  
/ LENGTH: 301  
/ TYPE: PRT  
/ ORGANISM: H. sapiens  
/ FEATURE:  
/ NAME/KEY: VARIANT  
/ LOCATION: (1)...(301)  
/ OTHER INFORMATION: TIM-3, allele 2  
US-10-188-012-31

Query Match 100.0%; Score 1587; DB 15; Length 301;  
Best Local Similarity 100.0%; Pred. No. 2.7e-151;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSLPFDVLLLLLLLTSSEVEYRAEYVGNAYLPCFYTPAAGNLVPCWGKACPV 60  
DB 1 MFSLPFDVLLLLLLLTSSEVEYRAEYVGNAYLPCFYTPAAGNLVPCWGKACPV 60  
QY 61 FECGNVVLRTDERDVNWTSRWYLNQDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
DB 61 FECGNVVLRTDERDVNWTSRWYLNQDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
QY 121 EKFNKLVIKPAKVTAPTLQRFPAAPRMLTTRGHGPAETQTLSLPDINTQISTLA 180  
DB 121 EKFNKLVIKPAKVTAPTLQRFPAAPRMLTTRGHGPAETQTLSLPDINTQISTLA 180  
QY 181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNLSLI 240  
DB 181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNLSLI 240

Db 181 NELSDSRDLANDLRSGATIRIGIYIGAGICAGLALALIFGALIPKWSHSEKIQNSLI 240  
QY 241 SLANLPPSGLANAAVEGIRSEENIYTIENYVEVEEPEEYCYSSROQPSQPGCRPFAM 300  
Db 241 SLANLPPSGLANAAVEGIRSEENIYTIENYVEVEEPEEYCYSSROQPSQPGCRPFAM 300  
QY 301 P 301  
Db 301 P 301

RESULT 4  
US-10-188-012-29  
; Sequence 29, Application US/10188012  
; Publication No. US20030124114A1  
; GENERAL INFORMATION:  
; APPLICANT: McInliffe, Jennifer Jones  
; APPLICANT: Umetsu, Dale T.  
; APPLICANT: Dekruyff, Rosemarie  
; APPLICANT: Kuchroo, Vijay  
; APPLICANT: Freeman, Gordon J.  
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
; FILE REFERENCE: STAN-235  
; CURRENT APPLICATION NUMBER: US/10/188,012  
; CURRENT FILING DATE: 2002-07-01  
; PRIOR APPLICATION NUMBER: 60/302,344  
; PRIOR FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 301  
; TYPE: PRT  
; ORGANISM: H. sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(301)  
; OTHER INFORMATION: TIM-3, allele 1  
US-10-188-012-29

Query Match 99.6%; Score 1581; DB 15; Length 301;  
Best Local Similarity 99.7%; Pred. No. 1.1e-150;  
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MFSLPDCVLLLLLLLITRSSEVEYRAEVGQNNYLPCFTYPAAPGNLVPVCMKGACPV 60  
Db 1 MFSLPDCVLLLLLLLITRSSEVEYRAEVGQNNYLPCFTYPAAPGNLVPVCMKGACPV 60  
QY 61 FECGNVLRTERDVNWTSRWYLNQDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
Db 61 FECGNVLRTERDVNWTSRWYLNQDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
QY 121 EKFNKLVIKPAKTPAPTLQDFTAAPPRMLTTRGHGPAETQTLGSLPDINTLQISTLA 180  
Db 121 EKFNKLVIKPAKTPAPTLQDFTAAPPRMLTTRGHGPAETQTLGSLPDINTLQISTLA 180  
QY 121 EKFNKLVIKPAKTPAPTLQDFTAAPPRMLTTRGHGPAETQTLGSLPDINTLQISTLA 180  
Db 121 EKFNKLVIKPAKTPAPTLQDFTAAPPRMLTTRGHGPAETQTLGSLPDINTLQISTLA 180  
QY 181 NELSDSRDLANDLRSGATIRIGIYIGAGICAGLALALIFGALIPKWSHSEKIQNSLI 240  
Db 181 NELSDSRDLANDLRSGATIRIGIYIGAGICAGLALALIFGALIPKWSHSEKIQNSLI 240  
QY 241 SLANLPPSGLANAAVEGIRSEENIYTIENYVEVEEPEEYCYSSROQPSQPGCRPFAM 300  
Db 241 SLANLPPSGLANAAVEGIRSEENIYTIENYVEVEEPEEYCYSSROQPSQPGCRPFAM 300  
QY 301 P 301  
Db 301 P 301

RESULT 5  
US-09-529-063-34  
; Sequence 34, Application US/09529063

; Patent No. US20020102542A1  
; GENERAL INFORMATION:  
; APPLICANT: FUKUSHIMA, DAIKICHI  
; APPLICANT: SHIBAYAMA, SHIRO  
; APPLICANT: TADA, HIDEAKI  
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF  
; FILE REFERENCE: THE BOTH  
; FILE REFERENCE: Q58769  
; CURRENT APPLICATION NUMBER: US/09/529,063  
; PRIOR APPLICATION NUMBER: PCT/JP98/04514  
; PRIOR FILING DATE: 1998-10-06  
; PRIOR APPLICATION NUMBER: JP 9-274674  
; PRIOR FILING DATE: 1997-10-07  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-063-34

Query Match 59.2%; Score 940; DB 10; Length 185;  
Best Local Similarity 99.4%; Pred. No. 1.3e-86;  
Matches 175; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MFSLPDCVLLLLLLLITRSSEVEYRAEVGQNNYLPCFTYPAAPGNLVPVCMKGACPV 60  
Db 1 MFSLPDCVLLLLLLLITRSSEVEYRAEVGQNNYLPCFTYPAAPGNLVPVCMKGACPV 60  
QY 61 FECGNVLRTERDVNWTSRWYLNQDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
Db 61 FECGNVLRTERDVNWTSRWYLNQDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
QY 121 EKFNKLVIKPAKTPAPTLQDFTAAPPRMLTTRGHGPAETQTLGSLPDINTLQISTLA 176  
Db 121 EKFNKLVIKPAKTPAPTLQDFTAAPPRMLTTRGHGPAETQTLGSLPDINTLQISTLA 176

RESULT 6  
US-10-414-378-34  
; Sequence 34, Application US/10414378  
; Publication No. US20030165981A1  
; GENERAL INFORMATION:  
; APPLICANT: FUKUSHIMA, DAIKICHI  
; APPLICANT: SHIBAYAMA, SHIRO  
; APPLICANT: TADA, HIDEAKI  
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF  
; FILE REFERENCE: Q58769  
; CURRENT APPLICATION NUMBER: US/10/414,378  
; PRIOR APPLICATION NUMBER: US/09/529,063  
; PRIOR FILING DATE: 2003-04-16  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: PCT/JP98/04514  
; PRIOR FILING DATE: 1998-10-06  
; PRIOR APPLICATION NUMBER: JP 9-274674  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-414-378-34

Query Match 59.2%; Score 940; DB 12; Length 185;  
Best Local Similarity 99.4%; Pred. No. 1.9e-86;  
Matches 175; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MFSLPDCVLLLLLLLITRSSEVEYRAEVGQNNYLPCFTYPAAPGNLVPVCMKGACPV 60  
Db 1 MFSLPDCVLLLLLLLITRSSEVEYRAEVGQNNYLPCFTYPAAPGNLVPVCMKGACPV 60

QY 61 FECGNVLTDERDNYWYTSRYWLNDFPKGDSVLTTEENVTLADSGIYCCRIQIPGIMND 120  
 DB 61 FECGNVLTDERDNYWYTSRYWLNDFPKGDSVLTTEENVTLADSGIYCCRIQIPGIMND 120  
 QY 121 EKFNKLKLVKPAKVTPTLPQDFTAPPRMLTTRGHGPAETOTLGSIPDINLTQI 176  
 DB 121 EKFNKLKLVKPAKVTPTLPQDFTAPPRMLTTRGHGPAETOTLGSIPDINLTQI 176

## RESULT 7

US-10-188-012-11  
 ; Sequence 11, Application US/10188012  
 ; Publication No. US20030124114A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McIntire, Jennifer Jones  
 ; APPLICANT: Umetau, Dale T.  
 ; APPLICANT: Dekruyff, Rosemarie  
 ; APPLICANT: Kuchroo, Vijay  
 ; APPLICANT: Freeman, Gordon J.  
 ; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
 ; FILE REFERENCE: STAN-235  
 ; CURRENT APPLICATION NUMBER: US/10/188,012  
 ; PRIOR FILING DATE: 2002-07-01  
 ; PRIOR APPLICATION NUMBER: 60/302,344  
 ; NUMBER OF SEQ ID NOS: 36  
 ; SOFTWARE: FastSeq for windows Version 4.0  
 ; SEQ ID NO 11  
 ; LENGTH: 281  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; FEATURE:  
 ; NAME/KEY: VARIANT  
 ; LOCATION: (1)...(281)  
 ; OTHER INFORMATION: TIM-3, C.D2 ES-HBA and DBA/2J allele  
 US-10-188-012-11

Query Match 55.8%; Score 885; DB 15; Length 281;  
 Best Local Similarity 64.0%; Pred. No. 1.2e-80;  
 Matches 187; Conservative 26; Mismatches 67; Indels 12; Gaps 4;  
 QY 1 MESHLPDCCVLLLLLLLTSSSEVEYRAEYGNATPCPTTPAAGNLYPVCGKACPV 60  
 DB 1 MESHLPDCCVLLLLLLLTSSSEVEYRAEYGNATPCPTTPAAGNLYPVCGKACPV 60  
 QY 61 FECGNVLTDERDNYWYTSRYWLNDFPKGDSVLTTEENVTLADSGIYCCRIQIPGIMN 119  
 DB 61 SGTNBLRTDERDNYWYTSRYWLNDFPKGDSVLTTEENVTLADSGIYCCRIQIPGIMN 120  
 QY 120 DEKNLKLKLVKPAKVTPTLPQDFTAPPRMLTTRGHGPAETOTLGSIPDINLTQI 179  
 DB 120 DEKNLKLKLVKPAKVTPTLPQDFTAPPRMLTTRGHGPAETOTLGSIPDINLTQI 179  
 QY 121 DKLELKLKLVKPAKVTPTLPQDFTAPPRMLTTRGHGPAETOTLGSIPDINLTQI 179  
 DB 121 DKLELKLKLVKPAKVTPTLPQDFTAPPRMLTTRGHGPAETOTLGSIPDINLTQI 179  
 QY 180 ANELRLSRRLANDRSGATIRIGIYAGICGGLALALIFGLIKWYSHSEKIQNLSL 239  
 DB 180 ANELRLSRRLANDRSGATIRIGIYAGICGGLALALIFGLIKWYSHSEKIQNLSL 239  
 QY 240 ISLANLPESGLANAAVEGIRSEENIYITEENVYEEVEEPNEYCYVSSROOPS 291  
 DB 240 ISLANLPESGLANAAVEGIRSEENIYITEENVYEEVEEPNEYCYVSSROOPS 291  
 QY 231 ITLANPPEGGLANAGAVRIRSEENIYITEENVYEEVEENNEYCYVNS-QQPS 281  
 DB 231 ITLANPPEGGLANAGAVRIRSEENIYITEENVYEEVEENNEYCYVNS-QQPS 281

## RESULT 8

US-10-252-131-10  
 ; Sequence 10, Application US/10252131  
 ; Publication No. US20030158399A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Levinson, Douglas A.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 ; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
 ; NUMBER OF SEQUENCES: 37

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036/2711

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/252,131  
 FILING DATE: 20-Sep-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/609,583  
 FILING DATE: 01-MAR-1996  
 APPLICATION NUMBER: US 08/487,748  
 FILING DATE: 07-JUN-1995  
 APPLICATION NUMBER: US 08/398,633  
 FILING DATE: 03-MAR-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7853-048  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-790-9090  
 TELEFAX: 212-869-8864  
 TELEX: 66141 PENNIE

## INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 281 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
 US-10-252-131-10

Query Match 55.6%; Score 883; DB 12; Length 281;  
 Best Local Similarity 63.7%; Pred. No. 1.9e-80;  
 Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;  
 QY 1 MESHLPDCCVLLLLLLLTSSSEVEYRAEYGNATPCPTTPAAGNLYPVCGKACPV 60  
 DB 1 MESHLPDCCVLLLLLLLTSSSEVEYRAEYGNATPCPTTPAAGNLYPVCGKACPV 60  
 QY 61 FECGNVLTDERDNYWYTSRYWLNDFPKGDSVLTTEENVTLADSGIYCCRIQIPGIMN 119  
 DB 61 SGTNBLRTDERDNYWYTSRYWLNDFPKGDSVLTTEENVTLADSGIYCCRIQIPGIMN 120  
 QY 120 DEKNLKLKLVKPAKVTPTLPQDFTAPPRMLTTRGHGPAETOTLGSIPDINLTQI 179  
 DB 120 DEKNLKLKLVKPAKVTPTLPQDFTAPPRMLTTRGHGPAETOTLGSIPDINLTQI 179  
 QY 121 DKLELKLKLVKPAKVTPTLPQDFTAPPRMLTTRGHGPAETOTLGSIPDINLTQI 179  
 DB 121 DKLELKLKLVKPAKVTPTLPQDFTAPPRMLTTRGHGPAETOTLGSIPDINLTQI 179  
 QY 180 ANELRLSRRLANDRSGATIRIGIYAGICGGLALALIFGLIKWYSHSEKIQNLSL 239  
 DB 180 ANELRLSRRLANDRSGATIRIGIYAGICGGLALALIFGLIKWYSHSEKIQNLSL 239  
 QY 240 ISLANLPESGLANAAVEGIRSEENIYITEENVYEEVEEPNEYCYVSSROOPS 291  
 DB 240 ISLANLPESGLANAAVEGIRSEENIYITEENVYEEVEEPNEYCYVSSROOPS 291  
 QY 231 ITLANPPEGGLANAGAVRIRSEENIYITEENVYEEVEENNEYCYVNS-QQPS 281  
 DB 231 ITLANPPEGGLANAGAVRIRSEENIYITEENVYEEVEENNEYCYVNS-QQPS 281

## RESULT 9

US-10-004-633-10  
 ; Sequence 10, Application US/10004633  
 ; Publication No. US20030069196A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Levinson, Douglas A.  
 ; APPLICANT: Lloyd, Clare M.  
 ; APPLICANT: McCarthy, Sean A.



TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
FILE OF INVENTION: DIAGNOSIS OF IMMUNE DISORDERS  
FILE REFERENCE: 7853-125  
CURRENT APPLICATION NUMBER: US/10/004,633  
PRIOR FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: 09/324,986  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583  
PRIOR FILING DATE: EARLIER FILING DATE: 1996-03-01  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/487,748  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/398,633  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-03-05  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 10  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-004-633-10

Query Match 55.6%; Score 883; DB 15; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.9e-80;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFDVCVLLLLLLLTSSSEVEYRAEVGNAYLPCFYTPAAGNLVPCWKGACPV 60  
1 MFSGTLNCVLLLLQLLARSLBEGYKVEGKNALPSCSYTLPTSGTLVPCWKGACPV 60  
DB 1 MFSGTLNCVLLLLQLLARSLBEGYKVEGKNALPSCSYTLPTSGTLVPCWKGACPV 60  
QY 61 FECCGVVLRTERDVNTVTSRYWNGDFRKGDSLTIENTVTLADSGIYCCRIQIPGIMN 119  
61 SQCTNELRTDERNVTVYKSSRYQKGLNKGDVSLIKNTVLDHGTCCRIQIPGIMN 120  
DB 61 SQCTNELRTDERNVTVYKSSRYQKGLNKGDVSLIKNTVLDHGTCCRIQIPGIMN 120  
QY 120 DEKNLKVIRPAKVTAPLQDFTAPFPMLTTRGHGPAETQISLPDINTQISTL 179  
121 DKKEKLDIKAAYTPAQTAGDSTSPRTLTERRNG-SETQTLVTLHNNNGTKISTW 179  
DB 121 DKKEKLDIKAAYTPAQTAGDSTSPRTLTERRNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELDRSLANDLDSGATIRIGIYGAGICAGLALIFGALIFKWTSHSKKXKIQNL 239  
180 ADE-----IKDSGETIRTAHIGVGSAGLTALIIIGVILKWTSCCKKLSSLSL 230  
DB 180 ADE-----IKDSGETIRTAHIGVGSAGLTALIIIGVILKWTSCCKKLSSLSL 230  
QY 240 ISLANLPSSGLANAVABGIRSEENIYTIENVEVEEPENEYYCYVSSRQOPS 291  
231 ITLANLPSSGLANAGAVRIRSEENIYTIENVEVEEPENEYYCYVNS-QOPS 281  
DB 231 ITLANLPSSGLANAGAVRIRSEENIYTIENVEVEEPENEYYCYVNS-QOPS 281

## RESULT 10

US-10-188-012-9  
Sequence 9, Application US/10188012  
Publication No. US20030124114A1  
GENERAL INFORMATION:  
APPLICANT: McIntire, Jennifer Jones  
APPLICANT: Umetsu, Dale T.  
APPLICANT: Dekruyff, Rosemarie  
APPLICANT: Kuchroo, Vijay  
APPLICANT: Freeman, Gordon J.  
TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
FILE REFERENCE: STAN-235  
CURRENT APPLICATION NUMBER: US/10/188,012  
CURRENT FILING DATE: 2002-07-01  
PRIOR APPLICATION NUMBER: 60/302,344  
PRIOR FILING DATE: 2001-06-29  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(281)  
OTHER INFORMATION: TIM-3 BALB/c allele

## US-10-188-012-9

Query Match 55.6%; Score 883; DB 15; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.9e-80;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFDVCVLLLLLLLTSSSEVEYRAEVGNAYLPCFYTPAAGNLVPCWKGACPV 60  
1 MFSGTLNCVLLLLQLLARSLBEGYKVEGKNALPSCSYTLPTSGTLVPCWKGACPV 60  
DB 1 MFSGTLNCVLLLLQLLARSLBEGYKVEGKNALPSCSYTLPTSGTLVPCWKGACPV 60  
QY 61 FECCGVVLRTERDVNTVTSRYWNGDFRKGDSLTIENTVTLADSGIYCCRIQIPGIMN 119  
61 SQCTNELRTDERNVTVYKSSRYQKGLNKGDVSLIKNTVLDHGTCCRIQIPGIMN 120  
DB 61 SQCTNELRTDERNVTVYKSSRYQKGLNKGDVSLIKNTVLDHGTCCRIQIPGIMN 120  
QY 120 DEKNLKVIRPAKVTAPLQDFTAPFPMLTTRGHGPAETQISLPDINTQISTL 179  
121 DKKEKLDIKAAYTPAQTAGDSTSPRTLTERRNG-SETQTLVTLHNNNGTKISTW 179  
DB 121 DKKEKLDIKAAYTPAQTAGDSTSPRTLTERRNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELDRSLANDLDSGATIRIGIYGAGICAGLALIFGALIFKWTSHSKKXKIQNL 239  
180 ADE-----IKDSGETIRTAHIGVGSAGLTALIIIGVILKWTSCCKKLSSLSL 230  
DB 180 ADE-----IKDSGETIRTAHIGVGSAGLTALIIIGVILKWTSCCKKLSSLSL 230  
QY 240 ISLANLPSSGLANAVABGIRSEENIYTIENVEVEEPENEYYCYVSSRQOPS 291  
231 ITLANLPSSGLANAGAVRIRSEENIYTIENVEVEEPENEYYCYVNS-QOPS 281  
DB 231 ITLANLPSSGLANAGAVRIRSEENIYTIENVEVEEPENEYYCYVNS-QOPS 281

## RESULT 11

US-09-966-546-22  
Sequence 22, Application US/09966546  
Patent No. US20020168716A1  
GENERAL INFORMATION:  
APPLICANT: Fernandes, Elma  
APPLICANT: Vernier, Corine  
APPLICANT: Shinkens, Richard A.  
TITLE OF INVENTION: No. US20020168716A1 Human Proteins and Polynucleotides Encoding  
FILE REFERENCE: Cura-46 (15966-546)  
CURRENT APPLICATION NUMBER: US/09/966,546  
CURRENT FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 09/544,511  
PRIOR FILING DATE: 2000-04-06  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 22  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-966-546-22

Query Match 45.2%; Score 718; DB 10; Length 142;  
Best Local Similarity 97.8%; Pred. No. 3e-64;  
Matches 132; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MFSLPFDVCVLLLLLLLTSSSEVEYRAEVGNAYLPCFYTPAAGNLVPCWKGACPV 60  
1 MFSGTLNCVLLLLQLLARSLBEGYKVEGKNALPSCSYTLPTSGTLVPCWKGACPV 60  
DB 1 MFSGTLNCVLLLLQLLARSLBEGYKVEGKNALPSCSYTLPTSGTLVPCWKGACPV 60  
QY 61 FECCGVVLRTERDVNTVTSRYWNGDFRKGDSLTIENTVTLADSGIYCCRIQIPGIMN 120  
61 FECCGVVLRTERDVNTVTSRYWNGDFRKGDSLTIENTVTLADSGIYCCRIQIPGIMN 120  
DB 61 FECCGVVLRTERDVNTVTSRYWNGDFRKGDSLTIENTVTLADSGIYCCRIQIPGIMN 120  
QY 121 EKFNKLVIRKPAKVT 135  
121 EKFNKLVIRKPAKVT 135  
DB 121 EKFNKLVIRKPAKVT 135

## RESULT 12

US-09-966-545-22  
Sequence 22, Application US/09966545  
Patent No. US20020172999A1  
GENERAL INFORMATION:  
APPLICANT: Fernandes, Elma

```
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US2002017999A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 142
; SEQ ID NO 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-545-22
```

```
Query Match      45.2%; Score 718; DB 10; Length 142;
Best Local Similarity 97.8%; Pred. No. 3e-64;
Matches 132; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 MFSLPFDVCVLLLLLTSSSEVEYRAEVGQNAVLPCEYTPAPAGNLVPCWGKACPV 60
        |||
Db      1 MFSLPFDVCVLLLLLTSSSEVEYRAEVGQNAVLPCEYTPAPAGNLVPCWGKACPV 60
        |||
QY      61 FECGNVVLRTDERDVNYWTSRYWLNDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120
        |||
Db      61 FECGNVVLRTDERDVNYWTSRYWLNDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120
        |||
QY      121 EKFNKLVIKPAKVT 135
        |||
Db      121 EKFNKLVIKPGEMT 135
        |||
```

```
RESULT 13
US-09-965-212-22
; Sequence 22, Application US/09965212
; Publication No. US20030003462A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20030003462A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/965,212
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: USSN 60/128,514
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-212-22
```

```
Query Match      45.2%; Score 718; DB 11; Length 142;
Best Local Similarity 97.8%; Pred. No. 3e-64;
Matches 132; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 MFSLPFDVCVLLLLLTSSSEVEYRAEVGQNAVLPCEYTPAPAGNLVPCWGKACPV 60
        |||
Db      1 MFSLPFDVCVLLLLLTSSSEVEYRAEVGQNAVLPCEYTPAPAGNLVPCWGKACPV 60
        |||
QY      61 FECGNVVLRTDERDVNYWTSRYWLNDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120
        |||
Db      61 FECGNVVLRTDERDVNYWTSRYWLNDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120
        |||
QY      121 EKFNKLVIKPAKVT 135
        |||
Db      121 EKFNKLVIKPAKVT 135
        |||
```

```
Db      121 EKFNKLVIKPGEMT 135
        |||
```

```
RESULT 14
US-10-189-940-22
; Sequence 22, Application US/10189940
; Publication No. US20030129613A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard
; APPLICANT: Anderson, David
; APPLICANT: Padigaru, Muraidhara
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh
; APPLICANT: Casman, Stacie
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: No. US20030129613A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: 15966-546 CIP
; CURRENT APPLICATION NUMBER: US/10/189,940
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/303,241
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/369,065
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/378,730
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 09/965,212
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,546
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/186,592
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 22
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-189-940-22
```

```
Query Match      45.2%; Score 718; DB 16; Length 142;
Best Local Similarity 97.8%; Pred. No. 3e-64;
Matches 132; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 MFSLPFDVCVLLLLLTSSSEVEYRAEVGQNAVLPCEYTPAPAGNLVPCWGKACPV 60
        |||
Db      1 MFSLPFDVCVLLLLLTSSSEVEYRAEVGQNAVLPCEYTPAPAGNLVPCWGKACPV 60
        |||
QY      61 FECGNVVLRTDERDVNYWTSRYWLNDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120
        |||
Db      61 FECGNVVLRTDERDVNYWTSRYWLNDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120
        |||
QY      121 EKFNKLVIKPAKVT 135
        |||
Db      121 EKFNKLVIKPGEMT 135
        |||

RESULT 15
US-10-189-940-146
; Sequence 146, Application US/10189940
; Publication No. US20030129613A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard
```

Query Match 45.2%; Score 718; DB 16; Length 142;  
Best Local Similarity 97.8%; Pred. No. 3e-64;  
Matches 132; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Search completed: November 22, 2003, 05:56:16  
Job time : 46.9948 secs

**THIS PAGE BLANK (USPTO)**

[illegible]



Db 355 LYATD 359

## RESULT 9

Q96D42 PRELIMINARY; PRT; 364 AA.  
 AC Q96D42; 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 OS Homo sapiens (Human).  
 OC Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Kidney;  
 RA Strausberg R.;  
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC013325; AAH13325.1; -.  
 DR Genew; HENC:17866; HAVCR1.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG\_Like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG\_1.  
 DR SMART; SM00409; IG\_1.  
 DR PROSITE; PS50835; IG\_Like; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 364 AA; 39249 MW; AA67C7DCTFAC91F1 CRC64;

Query Match 19.9%; Score 316.5; DB 4; Length 364;  
 Best Local Similarity 27.6%; Pred. No. 7.6e-22;  
 Matches 102; Conservative 42; Mismatches 111; Indels 115; Gaps 12;

QY 10 VLLLLLLLTSS--SEVERAEVGNAYLPCFYTPAAGNLYPVCMGKACVPECGNVV 67  
 DB 6 VILSLILHLADSVAGSVKVGEGAGPSVLPCHYS---GAVTSMCKMNGSGSLFCQNGI 61  
 QY 68 LRTDERDVNY-WTSRYWNGDFRKGDVSLTENTVTLADSGIYCCRIQIPGIMNDEKFNK 126  
 DB 62 VWTNGTHTVYRKDRTYKLLGDLSRDVSLLTENTAVDSGVYCCVEHKGEMFNKIVTS 121  
 QY 127 LVIRAKTTPAFTLORDTAAFPRLTT----- 154  
 DB 122 LEIVPCKVTTPIVTVPTVTRSTTVPTTTPMTVPTTTPPTTMSIPTTTVLT 181  
 QY 155 -----RGHGRATQTLGSLPDINLQI 176  
 DB 182 MVTSTTTSVPTTTSIPTTTSVPTTTSVTPVPMPLPRONHEPVATSP--SSPQPAETHP 239  
 QY 177 STLANELR---DSRIANDLRDSGAT-----RIGIYIG 206  
 DB 240 TTLQGAIRREPTSSPLSYTTDGNCTVTESSDGLMNNNOTQLFLEHSLTANTTGIT-- 297  
 QY 207 AGICGAG-LALALIFGALIFKWSHSEKIQNISLISLANLPPSGLANAVAGIRSEENITY 265  
 DB 238 AGVCISVIVLALLGVITIAKKYFKKE--VQQLS-VSFSLQIKALQNAVEKEVQAEEDNY 355  
 QY 266 TIEENVYEE 275  
 DB 356 -IENSLYATD 364

## RESULT 10

Q8VBMO PRELIMINARY; PRT; 305 AA.  
 AC Q8VBMO; 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE TIM2.  
 GN TIMD2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DA/2, and BALB/c; Tissue=Spleen;  
 RA McIntire J.J., Umetsu S.E., Akbari O., Potter M., Barsh G.S.,  
 RA Freeman G.J., Umetsu D.T., Dekruyf R.H.;  
 RT "Tapr", a major T cell regulatory locus that controls the development  
 of airway hyperactivity, cosegregates with variants in a novel gene  
 family."  
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF399828; AAJ35773.1; -.  
 DR EMBL; AF399827; AAJ35772.1; -.  
 DR MGI; MGI:2159681; Timd2.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG\_Like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG\_1.  
 DR SMART; SM00409; IG\_1.  
 DR PROSITE; PS50835; IG\_Like; 1.  
 SQ SEQUENCE 305 AA; 33519 MW; 68E69370FE23C9D0 CRC64;

Query Match 17.2%; Score 273.5; DB 11; Length 305;  
 Best Local Similarity 28.0%; Pred. No. 8.1e-18;  
 Matches 84; Conservative 41; Mismatches 134; Indels 41; Gaps 9;

QY 12 LLLLLLTSSSEVERAEVGNAYLPCFYTPAAGNLYPVCMGKACVPECGNVVLRD 71  
 DB 11 LILLPGAVESHNTAVQGLAGHPVLPCLYSTHL--GIVPMCKMGSGEGRHSICISLWTN 69  
 QY 72 ERDVNY-WTSRYWNGDFRKGDVSLTENTVTLADSGIYCCRIQIPGIMNDEKFNK 130  
 DB 70 GYVTHORNSRYQLKGNISEGNVSLTENTVTVGGDGPCCVVEIPGAFHFDVYMLE--VK 127  
 QY 131 PAKTPAFTLORDTAAFPRLTTRG--HGRATQTLGSLPDINLQI 171  
 DB 128 PEISTSEPT--RPAATGPRPTTSTRSTHVPSTIVSTSTSPATHTYKREATTTPDQ 185  
 QY 172 NLTOI-----STLANELDRSLANDLRDSGA-----TIRIGIYAGICAGIALA 216  
 DB 186 TTAETVETLPSTPADMNTVTSDDPMDNNEVAPQKPKQNLKGFVVGSI--AALLIL 244  
 QY 217 LIFGALIFKWSHSEKIQNISLISLANLPPSGLANAVAGIRSEENITYTIEENVYEE 276  
 DB 245 MLSTWVITRYVMKRSSESLSFVAFPIISKIGASPKVVERTRCEDQVYIIEPTPYPRE 304

## RESULT 11

Q8RI83 PRELIMINARY; PRT; 305 AA.  
 AC Q8RI83; 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical 33.5 Kda protein.  
 GN TIMD2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Liver;  
 RA Strausberg R.;  
 RL Submitted (Mar-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC025096; AAH25096.1; -.  
 DR MGI; MGI:2159681; Timd2.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG\_Like.





DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Hepatitis A virus receptor.  
 OS Cercopithecus aethiops (Green monkey) (Grivet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxID=9534;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=97368416; PubMed=9225030;  
 RX Ashida M., Hamada C.;  
 RT "Molecular cloning of the hepatitis A virus receptor from a simian  
 cell line."  
 RL J. Gen. Virol. 78:1565-1569(1997).  
 DR EMBL, D85851; BAA21556.1; -.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-1like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR InterPro; IPR002965; P-rich exten.  
 DR PRINTS; PR01217; PRICHEXTENSN.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR Receptor.  
 SQ SEQUENCE 460 AA; 49687 MW; 1D1A0C1D832139EA CRC64;

Query Match 15.7%; Score 248.5; DB 6; Length 460;  
 Best Local Similarity 22.5%; Pred. No. 3.6e-15;  
 Matches 103; Conservative 37; Mismatches 123; Indels 195; Gaps 13;

QY 4 HLPDPCVLLLLLLLRSEVEYRAVGNAYLPCFYTPAPAGNLVPCWKGACPFVEEC 63  
 DB 2 HLOVVISLILHLADPVDVAVDAGLSTLPKRN---GALISCMKNGTCSVESC 57  
 QY 64 GNVVLRTDERDVNY-WTSRYWLNQDFRKGDVSLTIENVTLASGIGCCRIQIPGIMNDEK 122  
 DB 58 PDGIWVTGHTVTKRETRYKLGNLNRDVSILTANRAVSDSGIYCCRVHSGWFMNPK 117  
 QY 123 FNLUKVIKPAKYT----- 135  
 DB 118 ITISLKIQPPRVTLIVRTVSTVPTTTTLPPTTLPPTTLPPTTLPPTTLP 177  
 QY 136 -----PAPTLQDRF--TAAPRM.LTRGHGPAETQTL----- 165  
 DB 178 TTTLPPTTTVPTTLPPTTLPPTTLPPTTLPPTTLPPTTLPPTTLPPTTLP 236  
 QY 166 -----GSLPDINLTQISTLANEL----- 183  
 DB 237 MTTLPPTTLPPTTLPPTTLPPTTLPPTTLPPTTLPPTTLPPTTLPPTTLP 296  
 QY 184 -----RDSRLANDLRDSGATI----- 199  
 DB 237 PPTPLPMONHEPVATSPSSPQPAETHPVTLGATRTQPTSPSPSYGTDGSDTVESDG 356  
 QY 200 -----RIGIYIGAGICAG-LALALIFGALIFKWSHKEKIONIS 238  
 DB 357 LMNNQOTQISPEHSQOMVNTTEGTY--AGVCISVLLAVLGVIAKKIFPKK--IQQLS 413  
 QY 239 LLSLANLPSPGLANAVAGIRSEENIYITIEENVYEEV 276  
 DB 414 -VSFSNHQKTLQNAVKKEVHAEDNIY-IENNLVAMNQ 449

RESULT 15

Q96H15 PRELIMINARY; PRT; 378 AA.  
 AC Q96H15;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Hypothetical protein.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=Brain;  
 RA Strausberg R.;  
 RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, BC008988; AAH08988.1; -.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-1like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 1.  
 DR Hypothetical protein.  
 SQ SEQUENCE 378 AA; 41550 MW; 758B9E90CB2BFE5D CRC64;

Query Match 15.5%; Score 245.5; DB 4; Length 378;  
 Best Local Similarity 36.2%; Pred. No. 5.4e-15;  
 Matches 63; Conservative 25; Mismatches 77; Indels 9; Gaps 4;

QY 15 LLLLRSEVEYRAVGNAYLPCFYTPAPAGNLVPCWKGACPFVECGNVLTATD-ER 73  
 DB 17 LYLFPVTSETVTEVLGHRVTLPCLYS-SWGHNSNSMCKGKQCPYSGCKEALITDGM 75  
 QY 74 DVNWTSTRYWLNQDFRKGDVSLTIENVTLASGIGCCRIQIPGIMNDEKFNLUKVIKPAK 133  
 DB 76 VTSRSKSAKYRLQGTIPRGDVSILTINPESDSGYCCRIEVPGWENDKINRLMLQKAS 135  
 QY 134 VT---PAPTLQDRFPAFP---RMLTRGHGPAETQTLGSLPDINLTQISTLA 180  
 DB 136 TTTTRATTTTRRTTTTPTTTRQMTTPALPPTTVVTTTPDITGTPTLQMTTIA 189

Search completed: November 22, 2003, 05:43:24  
 Job time : 61.4416 secs

**THIS PAGE BLANK (USPTO)**

	Matches	301,	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0
Qy		1	MFSLPFDVLLLLLLLLL	1	TRSEVEYRAEVGONAYLPCFYTPAPAGNLPVCMGKACPV				60	
Db		1	MFSLPFDVLLLLLLLLL	1	TRSEVEYRAEVGONAYLPCFYTPAPAGNLPVCMGKACPV				60	
Qy		61	FECCNVVLRTERDQVNWTSRYWLNCDPRKGVSLITINVTLADSGIYCCRIQIQGINMD						120	
Db		61	FECCNVVLRTERDQVNWTSRYWLNCDPRKGVSLITINVTLADSGIYCCRIQIQGINMD						120	
Qy		121	EKFNKLVIYKPAKYTPAPTLQRFDAFAEPRLMTTRGGHGAETQIIGSLPDIIMLTQISTLA						180	
Db		121	EKFNKLVIYKPAKYTPAPTLQRFDAFAEPRLMTTRGGHGAETQIIGSLPDIIMLTQISTLA						180	
Qy		181	NELRDSRLANDLRDSGATIRIGIYIAGICAGLALALIFGALIFKWSHSHKEXIQLNLSLI						240	
Db		181	NELRDSRLANDLRDSGATIRIGIYIAGICAGLALALIFGALIFKWSHSHKEXIQLNLSLI						240	
Qy		241	SLANLPBGLANAAVAGIRSEENIYTIENNVYEVEPNVYCYVSSRQPSQPLGCRFPM						300	
Db		241	SLANLPBGLANAAVAGIRSEENIYTIENNVYEVEPNVYCYVSSRQPSQPLGCRFPM						300	
Qy		301	SLANLPBGLANAAVAGIRSEENIYTIENNVYEVEPNVYCYVSSRQPSQPLGCRFPM						300	
Db		301	SLANLPBGLANAAVAGIRSEENIYTIENNVYEVEPNVYCYVSSRQPSQPLGCRFPM						300	
Qy		301	SLANLPBGLANAAVAGIRSEENIYTIENNVYEVEPNVYCYVSSRQPSQPLGCRFPM						300	
Db		301	SLANLPBGLANAAVAGIRSEENIYTIENNVYEVEPNVYCYVSSRQPSQPLGCRFPM						300	

RESULT 2	Q8TDD0	PRELIMINARY;	PRT;	301 AA.
ID	Q8TDD0			
AC	Q8TDD0;			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Hepatitis A virus cellular receptor 2.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Zhang W., Wan T., Li N., Cao X.;			
RT	"Novel human hepatitis A virus cellular receptor.":			
RL	Submitted (APR-2000) to the EMBL/Genbank/DBD databases.			
DR	EMBL; AF251707; AAM19100.1; -.			
DR	Genew; HGNC:18437; HAVCR2.			
DR	InterPro; IPR003599; IG.			
DR	InterPro; IPR007110; IG-like.			
DR	SMART; SM00409; IG; 1.			
DR	PROSITE; PS50835; IG_LIKE; 1.			
KW	Receptor.			
SO	SEQUENCE 301 AA; 3363 MW; 7DAEA13F01D3587B CRC64;			
Query Match 99.6%; Score 1581; DB 4; Length 301;				
Best Local Similarity 99.7%; Pred. No.1,4e-143;				
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0				
QY	1 MFSHLPDFCVLLLLLLLLTRSSSEVEYRAEVGONAVYPCFYTPAPGNLVPCWGKACPV 60			
DB	1 MFSHLPDFCVLLLLLLLLTRSSSEVEYRAEVGONAVYPCFYTPAPGNLVPCWGKACPV 60			
QY	61 PEGCGVVVLRTERDENVYVTSRWLNDFPKGQVSLTEENVTLADSGIYCCRIQIPIGMND 120			
DB	61 PEGCGVVVLRTERDENVYVTSRWLNDFPKGQVSLTEENVTLADSGIYCCRIQIPIGMND 120			
QY	121 EKFNKLTVIKPAKVTAPTLQORDFTAPFPRMLLTRGHGAETQTGLSPDINLTQISTLA 180			
DB	121 EKFNKLTVIKPAKVTAPTLQORDFTAPFPRMLLTRGHGAETQTGLSPDINLTQISTLA 180			
QY	181 NELRDSRLANDRDSGATIRIGIYGAGICAGALALLIGALLFKYSHSKEKIQNLSTL 240			
DB	181 NELRDSRLANDRDSGATIRIGIYGAGICAGALALLIGALLFKYSHSKEKIQNLSTL 240			

[illegible]

RESULT 3	
OSVIM0	
ID	PRELIMINARY;      PRT;      281 AA.
AC	OSVIM0;
DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	TM13.
GN	TM1D3.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_Taxid=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=DBA/2; TISSUE=Spleen;
RA	McIntire J.O., Umetsu S.E., Akbari O., Potter M., Barsh G.S.,
RA	Freeman G.J., Umetsu D.T., Dekruyff R.H.;
RT	"Tapir, a major T cell regulatory locus that controls the development
RT	of airway hyperactivity, cosegregates with variants in a novel gene
RT	family.";
RL	Submitted (JUL-2001) to the EMBL/genbank/DBDj databases.
DR	EMBL; AF299831; AA135776.1; .-
DR	MGI; MGI:2159682; TlmD3.
DR	InterPro; IPR003599; IG.
DR	InterPro; IPR007110; IG-like.
DR	InterPro; IPR000719; Prot_kinase.
DR	SMART; SM00409; IG; 1.
DR	PROSITE; PSS0835; IG_LIKE; 1.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
SO	SEQUENCE      281 AA;      30934 MW;      C0349E4BD0E5761D CRC64;

Query Match	55.8%	Score 985	DB 11	Length 281
Best Local Similarity	64.0%	Pred. No. 11e-76;		
Matches 187	Conservative 26	Mismatches 67	Indels 12	Gaps 4
QY	1 MFSLHPDCVLTLLILLTTRSSSEVEYRAVGONAYLPCEYTPAAPGMLVPVCWKGACPV 60			
Db	1 MFSGLTNCVLTLILLOLLARLSLENNAVPEVGNALVPCSYTLSTFGALVPMCWKGFPCPW 60			
QY	61 FECCNGVYLRRDEEDNYY--TSRYWYMGDPKRKDSVLTENTLADSGIYCRCIOIPGMN 11			
Db	61 SQCNELRLRDKENNVYOKSSRYQLKGDLINKDGVLLIKNYLDDHGYCCRIOEPGLMN 12			
QY	120 DEKENLKVIPAKVTPAETLORDFTAAPEPRMLITRGHPAETOQTGLSPDINTQTISTL 17			
Db	121 DKKELELDLIKAKVTPAQTAHQDSTPASPTLTTERNG-SETOTLVTLHNNGTKISTW 17			
QY	180 ANELRDRSRLANDRSDGATIRIGIYGAGICGGLALAIIFGALPKWYSHSKERIONSL 23			
Db	180 ADE-----IKDGETITRNRIHGVSASGLTALLIIGVLIINWSCKKKKLISLSTL 23			
QY	240 ISLANLPSPGLANAVAEGIRSEBNITYTIENYIEYEVEEPPNEYCYYSRSQQPS 291			
Db	231 ITLANLPFPGLANAGAVRIRISEBNITYTIENYIEVENSENSNEYCYVNS-QQPS 281			
 RESULT 4 Q8WM60 PRELIMINARY; PRT; 142 AA. Q8WM60; DT 01-MAR-2002 (TREMBLrel. 20, Created) DT 01-MAR-2002 (TREMBLrel. 20, last sequence update) DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)				

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 05:03:41 / Search time 58.4416 Seconds  
(without alignments)  
1329.085 Million cell updates/sec

Title: US-10-004-633-24

Perfect score: 1587

Sequence: 1 MESHLPFDCCVLLLLLLLTTR.....CYVSSROOPSPRIGRFAMP 301

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SPREMBL\_23.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1587	100.0	301	4 Q96K94	Q96K94 homo sapien
2	1581	99.6	301	4 Q8TDC0	Q8TDC0 homo sapien
3	885	55.8	281	11 Q8VIM0	Q8VIM0 mus musculu
4	718	45.2	142	4 Q8WV60	Q8WV60 homo sapien
5	336	21.2	307	11 Q54947	Q54947 rattus norv
6	333	21.0	282	11 Q8VIM1	Q8VIM1 mus musculu
7	323.5	20.4	305	11 Q8VIM2	Q8VIM2 mus musculu
8	322	20.3	359	4 Q43656	Q43656 homo sapien
9	316.5	19.9	364	4 Q96D42	Q96D42 homo sapien
10	273.5	17.2	305	11 Q8VIM0	Q8VIM0 mus musculu
11	263.5	16.6	305	11 Q8VIM0	Q8VIM0 mus musculu
12	256	16.1	451	6 Q95144	Q95144 mus musculu
13	248.5	15.7	362	11 Q8CIC7	Q8CIC7 mus musculu
14	248.5	15.7	460	6 Q18984	Q18984 cercopithec
15	245.5	15.3	378	4 Q96H15	Q96H15 homo sapien
16	243	15.3	474	6 Q46597	Q46597 cercopithec

17	241	15.2	478	6 Q46598	Q46598 cercopithec
18	137.5	8.7	258	4 Q9UMT2	Q9UMT2 homo sapien
19	135	8.5	276	4 Q95944	Q95944 homo sapien
20	133.5	8.4	258	4 Q9H563	Q9H563 homo sapien
21	132.5	8.3	270	4 Q9UMT1	Q9UMT1 homo sapien
22	131.5	8.3	336	4 Q8WV5	Q8WV5 homo sapien
23	131.5	8.3	352	11 Q91W66	Q91W66 mus musculu
24	131.5	8.3	365	11 Q9DBJ8	Q9DBJ8 mus musculu
25	131.5	8.3	523	4 Q00480	Q00480 homo sapien
26	131	8.3	276	4 Q9H562	Q9H562 homo sapien
27	130	8.2	319	6 Q9TU80	Q9TU80 homo sapien
28	130	8.2	344	11 Q9R067	Q9R067 canis fam11
29	130	8.2	358	11 Q9R067	Q9R067 rattus norv
30	128.5	8.1	270	4 Q9H564	Q9H564 mus musculu
31	125.5	7.9	319	6 Q9TU79	Q9TU79 sus scrofa
32	124.5	7.8	344	4 Q9UKV4	Q9UKV4 homo sapien
33	123	7.8	200	4 Q8WV7	Q8WV7 homo sapien
34	121	7.6	334	4 Q96AV7	Q96AV7 homo sapien
35	121	7.6	527	4 Q00475	Q00475 homo sapien
36	120	7.6	365	6 Q8WV3	Q8WV3 bos taurus
37	119.5	7.5	455	4 Q9UIR0	Q9UIR0 homo sapien
38	118	7.4	252	4 Q8WV16	Q8WV16 homo sapien
39	116.5	7.3	287	11 Q9D7B8	Q9D7B8 mus musculu
40	116	7.3	259	4 Q95532	Q95532 homo sapien
41	115	7.2	318	13 Q91664	Q91664 xenopus lae
42	114	7.2	789	11 Q8CJ59	Q8CJ59 mus musculu
43	112.5	7.1	430	4 Q8N4F1	Q8N4F1 homo sapien
44	111.5	7.0	319	11 Q922D5	Q922D5 mus musculu
45	111.5	7.0	529	4 P78408	P78408 homo sapien

## ALIGNMENTS

RESULT 1

Q96K94 PRELIMINARY; PRT; 301 AA.

AC Q96K94; 01-DEC-2001 (TRENBLREL. 19, Created)

DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)

DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)

DE Hypothetical protein Ftl14428 (Putative kidney injury molecule-3).

DE Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. OK NCBI\_taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RP TISSUE=Embryo;

RA Iisogai T., Ota T., Hayashi K., Sugiyama T., Ofuiki T., Suzuki Y., Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.; "NEBO human cDNA sequencing project."

RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

RU [2]

RP SEQUENCE FROM N.A.

RP TISSUE=Hepatoma;

RA Kuenen E.W., Ichimura T., Bonventre J.V.; "A homolog to human kidney injury molecule-1 is expressed in hepatoma cells."

RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

RU EMBL; AK077334; BAB55044.1; -

DR EMBL; AY069944; BAB55041.1; -

DR InterPro; IPR007110; IG-like.

DR PROSITE; PS50835; IG\_LIKE; 1.

KV Hypothetical protein\_

SQ SEQUENCE 301 AA; 33351 MW; 487787D9E40A4161 CRC64;

Query Match 100.0%; Score 1587; DB 4; Length 301;

Best Local Similarity 100.0%; Pred. No. 3.7e-144;

```

Qy 56 GACPVECGNVVLRTERDV---NYMTRWYMGDPFRKGVSLTIENVTLADSGIYCCR 111
Db 51 -----EPNSITLLAKSDVLFEDNYTKOXISYDSPYDDLVTITIKSLTARDAGTYVCA 104
Qy 112 IQIPGIMND-----EKENLKVIKPAKVTAPPTLQRPFTAAPPMLTTRGHGPAETOTL 165
Db 105 FPMSPPTNDTDKVDYEEYSTELIV-----NTDSESTIDILISGSTHSP---ETS 150
Qy 166 GSLP-DINLTQISTLANELRDSRLANDLADSGATIRIGIYGAGICAGLALALIFGALIF 224
Db 151 SEKPEDIDNLNCSSVFEIATPEPITDNVEDHTTV----- 185
Qy 225 KMYSHSKKIQNLISLISLANLPPSGIANAVAEGRSEENIYTIENV----- 271
Db 186 ---TYTSDSINTVSATS-----GESTTDETPPEPITDKEDHTVTDVSYTIVSTSSGIYV 237
Qy 272 --YEVEEPNEYCYVSSROQPSQPLGC 296
Db 238 TKSTTDADLYDTYNDNDTVFPSTVGC 264

```

Search completed: November 22, 2003, 05:41:06  
 Job time : 17.4811 secs

**THIS PAGE BLANK (USPTO)**

15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Butyrophilin precursor (BT).  
 GN BTNL1 OR BTN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_Taxid:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast; PubMed=8611614;  
 RX MEDLINE=96201966; Taylor M.R., Peterson J.A., Certani R.L., Couto J.R.;  
 RA "Cloning and sequence analysis of human butyrophilin reveals a  
 RT potential receptor function."  
 RL Biochim. Biophys. Acta 1305:1-4(1996).  
 CC -1- FUNCTION: MAY FUNCTION IN THE SECRETION OF MILK-FAT DROPLETS. IT  
 CC MAY ACT AS A SPECIFIC MEMBRANE-ASSOCIATED RECEPTOR FOR THE  
 CC ASSOCIATION OF CYTOLASMIC DROPLETS WITH THE APICAL PLASMA  
 CC MEMBRANE (BY SIMILARITY).  
 CC -1- SUBUNIT: SEEMS TO ASSOCIATE WITH XANTHINE DEHYDROGENASE/OXIDASE  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. BTN/MOG  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 CC EMBL; U39576; AAC50489.1; -.  
 CC PIR; S70587; S70587.  
 CC GeneW; HGNC:1135; BTNL1.  
 CC MIM; 601610; -.  
 CC DR GO; GO:0005887; C: integral to plasma membrane; TAS.  
 CC DR GO; GO:0004872; F: receptor activity; TAS.  
 CC DR InterPro; IPR001870; B302.  
 CC DR InterPro; IPR007110; IG-like.  
 CC DR InterPro; IPR003006; IG\_MHC.  
 CC DR InterPro; IPR003596; IG\_V.  
 CC DR InterPro; IPR006574; PRF.  
 CC DR InterPro; IPR003877; SPRY\_receptor.  
 CC DR Pfam; PF00047; IG\_1.  
 CC DR Pfam; PF00622; SPRY\_1.  
 CC DR SMART; SM00406; IGV\_1.  
 CC DR SMART; SM00589; PRY\_1.  
 CC DR SMART; SM00449; SPRY\_1.  
 CC DR PROSITE; PS50835; IG\_LIKE; 2.  
 CC Transmembrane; Glycoprotein; Immunoglobulin domain; signal; Repeat.  
 CC KW SIGNAL 1 26  
 CC FT CHAIN 27 526  
 CC FT DOMAIN 27 242  
 CC FT TRANSMEM 243 269  
 CC FT DOMAIN 270 526  
 CC FT DOMAIN 27 138  
 CC FT DOMAIN 148 234  
 CC FT CARBOHYD 55  
 CC FT CARBOHYD 215 215  
 CC SQ SEQUENCE 526 AA; 59004 MW; E9BCA0CFBDAF94D5 CRC64;

DB	FRKXVSPA	VLVHRRDREGEAEQMPYRGRATLVQGIKAGRAALNRGRVSDG	120
QY	107	IYCCRIQIPGIMNDEKFNKL	127
DB	121	EYTCFFREDGSYEALVHLKV	141
RESULT 15			
ID	HEMA_VACCT	STANDARD;	PRT; 315 AA.
AC	P16561;		
DT	01-AUG-1990	(Rel. 15, Created)	
DT	01-AUG-1990	(Rel. 15, Last sequence update)	
DT	15-JUL-1999	(Rel. 38, Last annotation update)	
DE	Hemagglutinin precursor.		
GN	HA OR A56R.		
OS	Vaccinia virus (strain Tian Tan).		
OC	Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;		
OC	Orthopoxvirus.		
OX	NCBI_Taxid=10253;		
RP	[1]		
SP	SEQUENCE FROM N.A.		
RX	MEDLINE=69328331; PubMed=2754392;		
RA	Dorogyan J., Zhiliang L., Qi J., Hao Y., Yunde H.;		
RT	"Vaccinia virus hemagglutinin. A novel member of the immunoglobulin		
RT	superfamily.";		
RL	J. Exp. Med. 170:571-576(1989).		
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.		
CC	-1- PPM: THE MAJOR 85 KDA GLYCOPROTEIN IS PHOSPHORYLATED AND SULFATED.		
CC	O-LINKED GLYCOSYLATION IS REQUIRED FOR HEMAGGLUTINATION AND		
CC	HEMAGGGLUTINATION ACTIVITIES OF INFECTED CELL MEMBRANES.		
CC	-1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
DR	EMBL; X15709; CAA33740.1; -.		
DR	PIR; J10108; HANZVT.		
DR	InterPro; IPR007110; Ig-like.		
DR	InterPro; IPR003599; Ig.		
DR	InterPro; IPR003006; Ig_MHC.		
DR	Pfam; PF00047; Ig; 1.		
DR	SMART; SMO0409; Ig; 1.		
DR	PROSITE; PSS0835; IG_LIKE; 1.		
KW	Hemagglutinin; Envelope protein; Late protein; Glycoprotein; Signal;		
KW	Transmembrane; Immunoglobulin domain; Phosphorylation; Sulfation.		
FT	CHAIN	1	16
FT	SIGNAL	17	315
FT	DOMAIN	17	279
FT	TRANSMEM	280	303
FT	DOMAIN	304	315
FT	DOMAIN	17	121
FT	DISULFD	34	103
FT	CARBOHYD	37	37
FT	CARBOHYD	69	69
FT	CARBOHYD	112	112
FT	CARBOHYD	161	161
FT	CARBOHYD	254	254
SEQUENCE	315 AA; 34772 MW; E049448640879FF4 CRC64;		
QY	Query Match	6.1%; Score 97.5; DB 1; Length 315;	
QY	Best Local Similarity	17.4%; Pred. No. 0.27;	
QY	Matches	57; Conservative 57; Mismatches 116; Indels 97; Gaps 12;	
DB	3 SHLPPDVLILLILTRSSVEYRAVGNAYLPCTYPAAGN-----LYPVCMKG	55	
DB	2 ARLP-----ILLILSLVSTSPQPSKIGDATALSC-----NRNNTNDVYVMSAYK	50	

```

FT FT isoform 5 and isoform 7).
FT FT /FTid=VSP 002920.
FT FT G-> EQWVTVSNVHNNSTASDNNRLG (in isoform
FT FT 8 and isoform 9).
FT FT /FTid=VSP 002921.
FT FT G-> VHGSEKGVLDQIWSLQDCDNEQVMTVNSDVH
FT FT NNSTASDNNRLG (in isoform 10 and isoform
FT FT 12).
FT FT /FTid=VSP 002922.
FT FT Missing (in isoform 4).
FT FT /FTid=VSP 002923.
SQ SEQUENCE 818 AA, 91736 MW, D1B39E2092B2152 CRC64,

Query Match 6.2%; Score 99; DB 1; Length 818;
Best Local Similarity 19.9%; Pred. No. 0.65;
Matches 59; Conservative 41; Mismatches 89; Indels 108; Gaps 12;

QY 87 DPKRQDVSLITNNVTLADSGIY--CQRIQIPG-INMDEK-----NKLVIKP----- 131
DB 241 DTSKRPASLTITKNVSMDSGLMISCAVENIVGEVOTSALVTFAPNITFIESPDPDHW 300
QY 132 -----AKYTPAPTLORDFTAAPR-----M 151
DB 301 CIPFVTKGNPKFTLQMFYEGALINSEYICTKIHVINOSEYVCCQLDNPHTLNNGAYTL 360
QY 152 LTRTGHGPAETQT-----LGSLEPDINLTQISTLANELRDSR-----LAN 190
DB 361 LAKNEYGEDEKRVDAHFMFSVPGDGGPIYDPPVYEYETTPNDLGGTNNNSQITSPDVSN 420
QY 191 DLKRDGATIRIGIYIGAGICAGIALALIFGALIFKMYSHSKKIKONTLSI----- 240
DB 421 KENEDSIYVVVVGIALVCTGLVIMLI-----ILKFRGSKFKMGKSPSVISNDDSDASP 476
QY 241 -----SLANLPPS--GLANAVAGIRSEENITYIENYVEVEPEHYCYVSSROOP 290
DB 477 LHHISNGSTPSSSEGGPDVAITGM-----TKLFIENP-QYFITTSQLKP 522

RESULT 13
MOG_RAT STANDARD; PRT; 245 AA.
ID 063145;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
PP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=93085763; PubMed=1453482;
RA Gardiner M.V., Amiguet P., Linington C., Mathieu J.-M.,
RT "Myelin/oligodendrocyte glycoprotein is a unique member of the
RT immunoglobulin superfamily."
RL J. Neurosci. Res. 33:177-187 (1992).
RN [2]
RP SEQUENCE OF 28-245 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93376728; PubMed=8367453;
RA Pham-Dinh D., Mattei M.-G., Nussbaum J.-L., Rousset G.,
RA Pontarotti P., Roeckel N., Mather I.H., Artzt K., Lindahl K.F.,
RA Dautigny A.,
RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the
RT immunoglobulin superfamily encoded within the major
RT histocompatibility complex."
RL Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994 (1993).
RN [3]
RP STRUCTURE BY NMR OF 62-82.
RX MEDLINE=97354172; PubMed=9210466;
RA Albouz-Abo S., Wilson J.C., Bernard C.C.A., von Itzstein M.;

```

```

RT "A conformational study of the human and rat encephalitogenic myelin
RT oligodendrocyte glycoprotein peptides 35-55."
RT Eur. J. Biochem. 246:59-70 (1997).
CC -1- FUNCTION: MINOR COMPONENT OF THE MYELIN SHEATH. MAY BE INVOLVED IN
CC COMPLETION AND/OR MAINTENANCE OF THE MYELIN SHEATH AND IN CELL-
CC CELL COMMUNICATION.
CC -1- SUBUNIT: MAY FORM HOMODIMERS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE CNS, WHERE IT IS
CC LOCALIZED ON THE SURFACE OF MYELIN AND OLIGODENDROCYTE CYTOPLASMIC
CC MEMBRANES.
CC -1- DEVELOPMENTAL STAGE: A PEAK OF EXPRESSION HAS BEEN OBSERVED
CC BETWEEN POSTNATAL DAYS 15 AND 25, COINCIDENT WITH THE PERIOD OF
CC ACTIVE MYELINATION.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. BTN/MOG
CC SUBFAMILY.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- CAUTION: DO NOT CONFUSE MYELIN-OLIGODENDROCYTE GLYCOPROTEIN (MOG)
CC WITH OLIGODENDROCYTE-MYELIN GLYCOPROTEIN (OMG).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M99485; AAA41628.1; -
DR EMBL, L21995; AAF74786.1; -
DR PIR, B47712; B47712.
DR InterPro: IPR007110; IG-1-like.
DR InterPro: IPR003066; IG_MHC.
DR InterPro: IPR003596; IG_V.
DR Pfam, PF00047; Ig_1.
DR SMART, SM00406; IGV_1.
DR PROSITE, PS50835; IG_LIKE; 1.
KW Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.
FT SIGNAL 1 27
FT CHAIN 28 245 MYELIN-OLIGODENDROCYTE GLYCOPROTEIN.
FT DOMAIN 28 155 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 156 176 POTENTIAL.
FT DOMAIN 177 208 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 209 229 POTENTIAL.
FT DOMAIN 230 245 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 30 139 IG-LIKE.
FT DISULFD 51 125 POTENTIAL.
FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 245 AA; 27881 MW; C97F8AD60D6A32B4 CRC64;

Query Match 6.2%; Score 98.5; DB 1; Length 245;
Best Local Similarity 27.2%; Pred. No. 0.16;
Matches 37; Conservative 18; Mismatches 66; Indels 15; Gaps 3;

QY 5 LPFDCVLLILLILTRSEVEY-----RAEYQNAVLPCEFTYPAAGNLVPCWGK 55
DB 9 LPSCLSLILLLOQSRSAVGQFVYIGPHPIRALVGDABALPCRSIPKNAITGMEVGYR 68
QY 56 GACPVEEGNAVLTDERDV---NYWTSRYWLNDGPKGVSLTIVNTVTLADSGIYCCR 111
DB 69 S--PFSRVVHLRYRNGKDDAQAPEYRGRTLEKESIGEGVALRIQVRFPSDEGGYTCF 126
QY 112 IQIPGIMNDEKFNKL 127
DB 127 FRDHSYQEAVALVKV 142

RESULT 14
BUTY_HUMAN STANDARD; PRT; 526 AA.
ID BUTY_HUMAN
AC Q13410;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

```





RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.U., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A., Smalls D.E.,  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences."  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1- FUNCTION: Probable immunoglobulin-like cell surface receptor.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=1;  
CC IsoId=Q9P1W8-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q9P1W8-2; Sequence=VSP\_007027;  
CC Note=No experimental confirmation available;  
CC Name=3;  
CC IsoId=Q9P1W8-3; Sequence=VSP\_007028;  
CC Note=No experimental confirmation available;  
CC -1- TISSUE SPECIFICITY: Detected in liver, and at very low levels in  
muscle, heart, lung, pancreas, kidney, placenta and skeletal  
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation-  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
CC EMBL, AB044624; BAA95692.1; -;  
CC EMBL, AL138804; CAC00474.1; -;  
CC EMBL, BC020629; AAH20629.1; ALT\_INIT.  
CC Genew; HGNC:15757; SIRPB2.  
CC MIM; 605466; -;  
DR GO; GO:0007267; P:cell-cell signaling; TAS.  
DR GO; GO:0007242; P:intracellular signaling cascade; TAS.  
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
DR InterPro; IPR003597; IG\_C1.  
DR InterPro; IPR003006; IG\_MHC.  
DR Pfam; PF00047; Igr\_3.  
DR SMART; SM00407; IGc1\_2.  
DR SMART; SM00406; IGV\_1.  
DR PROSITE; PS50835; IG\_LIKE; 3.  
DR PROSITE; PS00290; IG\_MHC; FALSE NEG.  
KV Repeat: Signal; Transmembrane; Immunoglobulin domain; Glycoprotein;  
KW Alternative splicing;  
FT SIGNAL 1 28  
FT CHAIN 29 387  
FT DOMAIN 29 360  
FT TRANSMEM 361 383  
FT DOMAIN 384 387  
FT DOMAIN 29 137  
FT DOMAIN 146 245  
FT DOMAIN 252 340  
FT DISULFID 168 119  
FT DISULFID 53 226  
FT DISULFID 271 329  
FT CARBOHYD 243 243  
FT CARBOHYD 268 268  
FT CARBOHYD 309 309  
FT CARBOHYD 317 317  
FT SIGNAL 1 28  
FT CHAIN 29 387  
FT DOMAIN 29 360  
FT TRANSMEM 361 383  
FT DOMAIN 384 387  
FT DOMAIN 29 137  
FT DOMAIN 146 245  
FT DOMAIN 252 340  
FT DISULFID 168 119  
FT DISULFID 53 226  
FT DISULFID 271 329  
FT CARBOHYD 243 243  
FT CARBOHYD 268 268  
FT CARBOHYD 309 309  
FT CARBOHYD 317 317

FT VARSPLIC 1 33 Missing (in isoform 2).  
FT FTID=VSP\_007027.  
FT VARSPLIC 144 360 Missing (in isoform 3).  
FT FTID=VSP\_007028.  
FT CONFLICT 263 263 A -> V (IN REF. 2).  
FT CONFLICT 286 286 L -> S (IN REF. 2).  
SQ SEQUENCE 387 AA; 42495 MW; PFF20C9F6E0B64B CRC64;  
Query Match 6.3%; Score 100.5; DB 1; Length 387;  
Best Local Similarity 25.5%; Pred. No. 0.19; Indels 59; Gaps 15;  
Matches 71; Conservative 37; Mismatches 111;  
QY 4 HLPDPCVLLLLLLLTTR-SSEVER- - - - -AEVQNAVLPCTYPAAGNVPYCW 53  
DB 9 HPGPFLLTLTLGLTEVAGEBELQIPKLLTVTKTATLHCTVSLP--VGPVLW 66  
QY 54 GKACPEVCEGVNVLRTDERDVNTSYWMLNGDFRK--DVSILTENVTLADSGIYC 110  
DB 67 FRGVGPGR-----LIYQKEGHF--PRVTTVSDLTKEKNMDFSISSITPADVGTYYC 119  
QY 111 RIQIPGINNDEK-----NLKVIKPAK---VTPAPTLQDFTAFPRMLTTRGHGPAE 161  
DB 120 VKFKGSPENVEFSGPCTEMALGAKPSAPVVLGPAATTPBHTVSF---TCSEHG--- 172  
QY 162 TQTLGSLPDINLTQISTLANELRDSRLANDRDSGATIRIGIYAGICAGIATLAIPGA 221  
DB 173 ----FSPRDITLTKMKK--GNELSD--FQTNVDPICQSAVYSRSTARV----- 213  
QY 222 LIFKWSHSHKEKIONLSLISLANLPPSGLANAVAGIR 259  
DB 214 VLDPWDVRS-OVICEVAHVTLQGDPLRGTA--LSEAIR 249  
RESULT 12  
TRKB\_CHICK  
ID TRKB\_CHICK STANDARD; PRT; 818 AA.  
AC Q91987; Q91010;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE BDNF/NF-3 growth factor receptor precursor (EC 2.7.1.112) (TrkB  
DE tyrosine kinase) (TrkB-B).  
GN TRKB.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 10).  
RC TISSUE=Brain;  
RX MEDLINE=95047511; PubMed=7959025;  
RA Vin N., Erdmann K., Heumann R.;  
RT "Cloning and sequence analysis of a cDNA encoding a novel truncated  
RT form of the chicken TrkB receptor.";  
RL Gene 149:383-384(1994).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=94116452; PubMed=8287802;  
RA Dechant G., Biffo S., Okazawa H., Kolbeck R., Potgiesser J.,  
RA Barde Y.A.;  
RT "Expression and binding characteristics of the BDNF receptor chick  
RT trkB.";  
RL Development 119:545-558(1993).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE FROM N.A.  
RP (ISOFORMS 2; 3; 4; 5; 6; 7; 8; 9; 10; 11 AND 12).  
RX MEDLINE=96370546; PubMed=8774442;  
RA Garner A.S., Mengay H.J., Boeshore K.L., Xie X.Y., Voci J.M.,  
RA Johnson J.E., Lange T.H.;  
RT "Expression of TrkB receptor isoforms in the developing avian visual  
RT system.";  
RL J. Neurosci. 16:1740-1752(1996).

CC - TISSUE SPECIFICITY: Ubiquitous.  
CC - SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
CC - SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC - DATABASE: NAME=PROW; NOTE=PROW 1:74-77(2000);  
WWW="http://www.ncbi.nlm.nih.gov/prow/guide/204270028.g.htm"  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; X80038; CAA56342.1; -  
DR EMBL; AF058448; AAC23797.1; -  
DR EMBL; BC003091; AAH03091.1; -  
DR EMBL; AF044968; AAC82348.1; -  
DR EMBL; AF044962; AAC82348.1; JOINED.  
DR EMBL; AF044963; AAC82348.1; JOINED.  
DR EMBL; AF044964; AAC82348.1; JOINED.  
DR EMBL; AF044966; AAC82348.1; JOINED.  
DR EMBL; AF044967; AAC82348.1; JOINED.  
DR EMBL; AF050154; AAC02503.1; -  
DR PIR; I68093; I68093.  
DR Genew; HGNC:9707; PVRL2.  
DR MIM; 600798; -  
DR GO; GO:0005886; C:plasma membrane; TAS.  
DR GO; GO:0015026; F:coreceptor activity; TAS.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig\_1.  
DR SMART; SM00409; Ig\_1.  
DR PROSITE; PS00835; IG\_LIKE; 3.  
DR Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; Signal;  
KW Repeat; Alternative splicing.  
FT SIGNAL 1 31  
FT CHAIN 32 538  
FT DOMAIN 32 360  
FT TRANSMEM 361 381  
FT DOMAIN 382 538  
FT DOMAIN 32 156  
FT DOMAIN 162 256  
FT DOMAIN 261 345  
FT DISULFID 54 140  
FT DISULFID 183 238  
FT DISULFID 283 329  
FT CARBOHYD 137 137  
FT CARBOHYD 324 324  
FT VARSPLIC 351 479  
SO SEQUENCE 538 AA; 57742 MW; 3ABF83E92F6F624 CRC64;  
Query Match 6.44; Score 101.5; DB 1; Length 538;  
Best Local Similarity 25.64; Pred. No. 0.23;  
Matches 57; Conservative 29; Mismatches 96; Indels 41; Gaps 13;

QY 11 LLLLLLLTSSRV-----EYAEVQGNAYLPCTTPAPRGVLP-VCKGKACVFEK 63  
DB 21 LLLLLLLETAQDVQVLPDEVRGQGGTVELPCHLLPVPGLYISLVTVQRDAPR-NH 79  
QY 64 GNVVL-----RTDERDVNVTSTRYWLNGFRRK--GVSILTIENVTLADSGIYC 109

DB 80 QNVAHFHKMGSPSPSPGSEERLSFVSAKOSTGQDTAEALQDATALHGLTVEDEGNYT 139  
QY 110 CRI-----QIPGIMNDEKFNILKVPKAYTPAP-TLQDFTAAPFPMLTRGHGPAE 161  
DB 140 CEFATPFGKSVRGMT-----WLRVAKRPNQAEQKVFQSDPFTV-ALCISKSGRPPAR 193  
QY 162 TQTLGSLP-DINLQIIS-TLANELR-DSR--LANDLRSGNTI 199  
DB 194 ISWSSLLDWEAKETQVSGTLGTVVTSRFTLVPSGRADGVTV 236  
RESULT 11  
ID SRB2\_HUMAN STANDARD; PRT; 387 AA.  
AC Q9PIW8; Q9WMA5; Q9NOK8;  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Signal-regulatory protein beta-2 precursor (SIRP-beta-2) (SIRP-b2).  
GN SIRPb2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC TISSUE=Placenta;  
RA MEDLINE=21036165; PubMed=1185750;  
RA Ichigori Y., Matsuda S., Machida K., Oshima K., Iwamoto T.,  
RA Yamaki K., Hayakawa T., Hanaguchi M.;  
RT "Molecular cloning of a novel human gene (SIRP-B2) which encodes a new  
RT member of the SIRP/SHPs-1 protein family.";  
RT J. Hum. Genet. 45:378-382 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Scavrides G., Almeida J.P., Babbage A.K., Bagnuley C.L.,  
RA Bailey O.P., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,  
RA Beasley J.C., Bird C.P., Blakey S.E., Bridgeman A., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grafham D.V., Griffiths M.N.D., Griffiths J.L., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle B., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Leharbain M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConnachie L.J., McEay J., McMuray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., McKerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,  
RA Skuce C.D., Smith M.N., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas A.D., Thorpe A.,  
RA Tracey A., Tsimas A.C., Vaundin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whitcaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RT Nature 414:865-871 (2001).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 3).  
RC TISSUE=Lung;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

OF THEIR LENGTH, BUT IS PRESENT ON NEW AND GROWING AXONS DERIVED FROM THE RGCS AT THE RETINAL MARGIN. REMAINS ON ADULT RGCS ONLY AT CELL-CELL CONTACT SITES AND IS CONTINUOUSLY FOUND IN THE RETINAL AXON TERMINAL ARBOR LAYERS OF THE ADULT TECTUM.

-1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.

-2- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: L25056; AAC38015.2; -

DR PIR: I50478; I50478.

DR HSP: Q13740; I10C.

DR InterPro: IPR007110; Ig-1-like.

DR InterPro: IPR003006; Ig\_MHC.

DR InterPro: IPR003596; Ig\_V.

DR Pfam: PF00047; Ig\_3.

DR SMART: SM00406; IGV, 1.

DR PROSITE: PS00835; IG\_LIKE, 4.

DR PROSITE: PS00290; IG\_MHC, 1.

KW Signal; Developmental protein; Cell adhesion; Immunoglobulin domain; Repeat; Glycoprotein; Transmembrane.

KW Repeat; Glycoprotein; Transmembrane.

FT CHAIN 1 22

FT SIGNAL 1 22

FT DOMAIN 23 555

FT TRANSMEM 23 499

FT DOMAIN 500 520

FT DOMAIN 521 555

FT DOMAIN 131 127

FT DOMAIN 131 229

FT DOMAIN 229 323

FT DOMAIN 319 397

FT DOMAIN 406 484

FT DISULFID 154 217

FT DISULFID 154 217

FT DISULFID 263 306

FT DISULFID 426 470

FT CARBOHYD 92 92

FT CARBOHYD 171 171

FT CARBOHYD 350 350

FT CARBOHYD 441 441

FT CARBOHYD 465 465

FT SEQUENCE 555 AA; 60371 MW; 5448014F00BFF68 CRC64;

Query Match 6.5%; Score 102.5; DB 1; Length 555;

Best Local Similarity 23.4%; Pred. No. 0.2;

Matches 44; Conservative 26; Mismatches 75; Indels 43; Gaps 7;

QY 78 WTSRWLNDPRKGVSLTEENTLADSGIYCCRIQIPGIMNDEKNTL----- 125

DB 359 WTK-----DNKKLKLDPFSKLTYSADGLVYCDVSIIEIKSLFELVVEGIPKITSLT 412

QY 126 -----KLVIKPAKVPAPLPORDFTAAPFRMLTTRGHGPAEQTLGSPDINTLQI 176

DB 413 KRRSDGKKKVLTCESGSPKPDVQSVNGTDEVSYNNKA---TYKLTVPSSKNLTIVS 469

QY 177 STLANEL-----RDSRIANDLRDGS--ATIRIGIYAGICAGLALIFGALIFK---- 225

DB 470 CLVTKLGLGDTKEISVFSQKNGDGTGAQAVIIGVIGLVLAAL-VGLLIYWIYIKKTRQG 528

QY 226 -WYSHSK 232

DB 529 SWKTGEKE 536

RESULT 10

PIR2\_HUMAN STANDARD; PRT; 538 AA.

AC Q92692; Q75455; Q96J29;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Poliovirus receptor related protein 2 precursor (herpes virus entry mediator B) (HvB) (Nectin 2) (CD112 antigen).

GN PVR2 OR PRR2 OR HVB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_Taxid:9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM DELTA).

RX MEDLINE=95347610; PubMed=7622062;

RA Eberle F., Dubreuil P., Mattei M.-G., Devillard E., Lopez M.; "The human PRR2 gene, related to the human poliovirus receptor gene (PVR), is the true homolog of the murine MPH gene"; Gene 159:267-272(1995).

RL

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM ALPHA).

RX MEDLINE=968321161; PubMed=9657005;

RA Warner M.S., Geraghty R.J., Martinez W.M., Montgomery R.I., Whitebeck J.C., Xu R., Eisenberg R.J., Cohen G.H., Spear P.G.; "A cell surface protein with herpesvirus entry activity (HvB) confers susceptibility to infection by mutants of herpes simplex virus type 1, herpes simplex virus type 2, and pseudorabies virus"; Virology 246:179-189(1998).

RL

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM ALPHA).

RC TISSUE=Brain;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Scheefter C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W., Villalón D.K., Muzny D.M., Sodergren E.U., Lu X., Gibbs R.A., Fahey J., Heltan E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M., Butterfield J.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RL

RN [4]

RP SEQUENCE OF 31-538 FROM N.A.

RT Yoshitua K., Murray J.C.; "A transcriptional map in the region of 19q13 derived using direct sequencing and exon trapping"; Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.

RL

RN [5]

RP SEQUENCE OF 449-538 FROM N.A.

RX MEDLINE=99449047; PubMed=10520737;

RA Freitas E.M., Zhang W.J., Lalonde J.P., Tay G.K., Gaudieri S., Ashworth L.K., Van Bockmeier F.M., Dawkins R.L.; "Sequencing of 42kb of the APO E-C2 gene cluster reveals a new gene: PERC1"; DNA Seq. 9:89-101(1998).

RL

CC -1- FUNCTION: RECEPTOR FOR ALPHAHERPEVIRUS (HSV-1, HSV-2 AND PSEUDORABIES VIRUS) ENTRY INTO CELLS.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=2;

CC Name=Delta;

CC Name=Alpha;

CC IsoId=Q92692-1; Sequence=Displayed;

CC IsoId=Q92692-2; Sequence=VSP\_002628, VSP\_002629;

FT DOMAIN 446 540 FIBROECTIN TYPE-III 1.  
 FT DOMAIN 543 639 FIBROECTIN TYPE-III 2.  
 FT DOMAIN 643 744 FIBROECTIN TYPE-III 3.  
 FT DOMAIN 839 1118 PROTEIN KINASE.  
 FT NP BIND 845 853 ATP (BY SIMILARITY).  
 FT BINDING 870 870 ATP (BY SIMILARITY).  
 FT ACT SITE 979 979 BY SIMILARITY.  
 FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 503 503 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 596 596 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT MOD RES 1007 1007 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 SQ SEQUENCE 1138 AA; 125089 MW; 38428B3678C58A1 CRC64;  
 Query Match 6.8%; Score 108; DB 1; Length 1138;  
 Best Local Similarity 23.6%; Pred. No. 0.15;  
 Matches 47; Conservative 32; Mismatches 46; Gaps 8;  
 QY 17 LITRSEVEYRAEYVGNAYLPCEYTPAPGNTVPCWGKACPFVPCGNAVLTDERDVN 76  
 DB 353 IINMASELEFNFETWPR--INC---AAAGNPPV---RGSIELRKPDDGLVLTSTVAIV- 402  
 QY 77 WYTSRWMLNGDRKGVSLTIENVTLADSGIYCCRIQIIMNDEKENLKVYKPAKTP 136  
 DB 403 -----EPEKTTAEFEVPRVLADSGFWEKCRSTSGGQDSRRFKVNV---KVP 447  
 QY 137 APTLORDFTAAPFRLTLTR-----GHGPAETQTLGSLPDLMLTOISTLANEL 183  
 DB 448 VP-----LAAPRLITKSRQQLVSPVFSGSDGISTVRLHYRQDSTMWSTIVDP 500  
 QY 184 RDSRLANDLR-DSGATIRI 201  
 DB 501 SENVTLMNLRPKGYSVRV 519  
 RESULT 8  
 HEMA\_RACVI STANDARD; PRT; 310 AA.  
 ID HEMA\_RACVI 000716;  
 AC 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Hemagglutinin precursor.  
 GN HA.  
 OS Raccoon poxvirus (RCN).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Orthopoxvirus.  
 OX NCBI\_TaxID=10256;  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=92410621; PubMed=1529542;  
 RA Cavallaro K.F., Esposito J.J.;  
 RT "Sequences of the raccoon poxvirus hemagglutinin protein.",  
 RL virology 190:434-439(1992).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- PTM: MAY BE PHOSPHORYLATED AND SULFATED. O-LINKED GLYCOSYLATION  
 CC MAY BE REQUIRED FOR HEMAGGLUTINATION AND HEMADSORPTION ACTIVITIES  
 CC OF INFECTED CELL MEMBRANES.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC EMBL; M94169; AAA47231.1; -.  
 DR PIR; A43381; HNVZRA.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003599; Ig.

DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig\_1.  
 DR SMART; SM00409; Ig\_1.  
 DR PROSITE; PS50835; Ig\_LIKE; 1.  
 KW Hemagglutinin; Envelope protein; Late protein; Glycoprotein; Signal;  
 KW Transmembrane; Immunoglobulin domain.  
 FT SIGNAL 1 16 POTENTIAL.  
 FT CHAIN 17 310 HEMAGGLUTININ.  
 FT DOMAIN 17 275 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 276 300 POTENTIAL.  
 FT DOMAIN 301 310 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 19 121 IG-LIKE V-TYPE.  
 FT DISULFID 36 105 POTENTIAL.  
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 310 AA; 34189 MW; 7C25F5C19C7B11B9 CRC64;  
 Query Match 6.7%; Score 107; DB 1; Length 310;  
 Best Local Similarity 23.2%; Pred. No. 0.038;  
 Matches 53; Conservative 36; Mismatches 81; Indels 58; Gaps 10;  
 QY 10 VILLILLILLITRSEVEYRAEYVGNAYLPCEYTPAPGNTVPCWGKACPFVPCGNAVLR 69  
 DB 8 ILLLSIVYTKKHPQISKGLGDDATLSC-NRNNTHGYLWSSWK-----KPSIILL 60  
 QY 70 TDERDVNWYTSRWMLNGDRKGVSLTIENVTLADSGIYCCRIQIIPGIM 118  
 DB 61 AAKNDVYVF-----DDVTADKVSVDSPYDPLATITIKSLTSDAGYICAFITTS-T 112  
 QY 119 ND-----EKENLKVYKPAKT-----PAPTLORDFTAAPFRLTLTRGHGPAET-QTL 165  
 DB 113 NDTKIDVEEYFIDLVNPNANVTIDALISGSTOODIISH-----TERQHSDDTICTS 167  
 QY 166 GSLPDINLTQISTLANELRDS-----RLANDLDSGATI 199  
 DB 168 ESTQISTSESTSSQISTSESTSSYGVEDDTQVNVTTDTTDSPTI 215  
 RESULT 9  
 C166\_CARAU STANDARD; PRT; 555 AA.  
 ID C166\_CARAU 090304;  
 AC 01-NOV-1997 (Rel. 35, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE CD166 antigen homolog precursor (Neurolin) (DM-GRASP homolog).  
 OS Carassius auratus (Goldfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Carassius.  
 OX NCBI\_TaxID=7957;  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Retina;  
 RC laesing U.;  
 RL submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 CC SEQUENCE OF 30-552 FROM N.A., AND PARTIAL SEQUENCE.  
 RP TISSUE=Retina;  
 RX MEDLINE=94299040; PubMed=8026643;  
 RA laesing U., Giordano S., Stecher B., Lottspeich F., Stuermer C.A.O.;  
 RT "Molecular characterization of fish neurolin: a growth-associated  
 RT cell surface protein and member of the immunoglobulin superfamily in  
 RT the fish retinotectal system with similarities to chick protein  
 RT DM-GRASP/SC-1/BEN.";  
 RL Differentiation 56:21-29(1994).  
 CC -1- FUNCTION: CELL ADHESION MOLECULE. INVOLVED IN NEURITE EXTENSION BY  
 CC NEURONS VIA HETEROPHILIC AND HOMOPHILIC INTERACTIONS.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: PRESENT ON ALL RETINAL GANGLION CELLS (RGCS)  
 CC AND THEIR AXONS (IN EMBRYO). ABSENT FROM MATURE AXONS ALONG MOST



RA Moritz R.L., Ji H., Heath J.K., White S.J., Welt S., Old L.J.,  
 RA Simpson R.J.,  
 RT "Characterization of posttranslational modifications of human A33  
 RT antigen, a novel palmitoylated surface glycoprotein of human  
 RT gastrointestinal epithelium.";  
 RL Biochem. Biophys. Res. Commun. 236:682-686 (1997).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN CELL-CELL RECOGNITION AND SIGNALING.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN NORMAL GASTROINTESTINAL  
 CC EPITHELIUM AND IN 95% OF COLON CANCERS.  
 CC -1- PTM: N-GLYCOSYLATED, CONTAINS APPROXIMATELY 8 KDA OF N-LINKED  
 CC CARBOHYDRATE.  
 CC -1- PTM: PALMITOYLATED.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain;  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U79725; AAC0957.1; -  
 CC Genew; HGNC:4445; GPA33.  
 DR MIM; 602171; -  
 DR GO; GO:0005888; C:proteoglycan integral to plasma membrane; TAS.  
 DR GO; GO:0004872; F:receptor activity; TAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig\_2.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PSS00335; IG LIKE; 2.  
 DR Immunoglobulin domain, lipoprotein; Palmitate; Glycoprotein;  
 KM Transmembrane; Signal; Antigen.  
 FT SIGNAL 1 21  
 FT CHAIN 22 319  
 FT DOMAIN 22 235  
 FT TRANSMEM 236 256  
 FT DOMAIN 257 319  
 FT DOMAIN 22 134  
 FT DOMAIN 140 227  
 FT DOMAIN 258 261  
 FT DISULFID 43 117  
 FT DISULFID 146 232  
 FT CARBOHYD 112 211  
 FT CARBOHYD 200 223  
 FT CARBOHYD 223 223  
 SQ SEQUENCE 319 AA; 35632 MW; 9BFC7AFAF45C2408E CRC64;  
 Query Match 7.1%; Score 112.5; DB 1; Length 319;  
 Best Local Similarity 22.7%; Pred. No. 0.013;  
 Matches 67; Conservative 47; Mismatches 108; Indels 73; Gaps 15;  
 QY 27 RAEVGNAYLPCEYTPAPGNTLVPCMGKACPFECGVNLTDERDV-----NY-- 77  
 DB 32 RASQSGSVTLPCYHSTSSREGLIOMDK-----LLLHTRRVYIWPSSNNYIH 81  
 QY 78 ---WTSRYLWNGDFRKGDVSLTENTVTLADSGYCCRIQPGIMDEKFNLK-----LV 128  
 DB 82 GELYKRVASISNNAEODASITTDQTLMDNGTYECSV--SI MSDLEGNTSRVLLVLI 138  
 QY 129 IKPAKTPAPTLORDPTAAPPRLTTRGH-----GPARTQTL 165  
 DB 139 VPPSK--PEGIGEGITIGNNIOLTCQSGEGSPTPOYSWKRYIINOQPLAOPASGQPV 196  
 QY 166 GSLPDIINTL---QISTLANELDSRLAN-DLDSGATIRIGIYGAGICAGIALAL-IF 219  
 DB 197 -SLKNISTDPSGYICTSSNE-EGTQFCNTITAVRSPSMVALYV--GIAVGVALLI 252

QY 220 GALLIKWTSXSEKQIQLSLISLANL-----PPSLANAVAEGIRSENIYITEE 269  
 DB 253 GIIIVCCCCRGKDWDTEDEKEDAPRREAYSEPPQLRELSRB--REEDDYROEE 305  
 RESULT 6  
 ID PVR2\_MOUSE STANDARD; PRT; 530 AA.  
 AC P32507; 062096;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Poliovirus receptor related protein 2 precursor (Murine herpesvirus  
 GN entry protein B) (mVeb) (Nectin 2) (Poliovirus receptor homolog).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA).  
 RX MEDLINE=92219365; PubMed=1560525;  
 RA Morrison M.E., Racanelli V.R.;  
 RT "Molecular cloning and expression of a murine homolog of the human  
 RT poliovirus receptor gene";  
 RL J. Virol. 66:2807-2813 (1992).  
 RN (2)  
 RP SEQUENCE FROM N.A. (ISOFORM BETA).  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=94179228; PubMed=8132569;  
 RA Aoki J., Koike S., Ise I., Sato-Yoshida Y., Nomoto A.;  
 RT "Amino acid residues on human poliovirus receptor involved in  
 RT interaction with poliovirus";  
 RL J. Biol. Chem. 269:8431-8438 (1994).  
 RN (3)  
 RP CHARACTERIZATION.  
 RX MEDLINE=99214397; PubMed=10196354;  
 RA Shukla D., Rowe C.L., Dong Y., Racanelli V.R., Spear P.G.;  
 RT "The murine homolog (Mpr) of human herpesvirus entry protein B (Hvrb)  
 RT mediates entry of pseudorabies virus but not herpes simplex virus  
 RT types 1 and 2.";  
 RL J. Virol. 73:4493-4497 (1999).  
 CC -1- FUNCTION: RECEPTOR FOR ALPHAHERPESVIRUS (SUCH AS MURINE HSV) ENTRY  
 CC INTO CELLS.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=Beta;  
 CC IsoId=P32507-1; Sequence=Displayed;  
 CC Name=Alpha;  
 CC IsoId=P32507-2; Sequence=VSP 002630, VSP 002631;  
 CC -1- TISSUE SPECIFICITY: BRAIN, SPINAL CORD, SPLEEN, KIDNEY, HEART AND  
 CC LIVER.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M80206; AAA39734.1; -  
 CC EMBL; D26107; BA05103.1; -  
 DR PIR; A38211; HLMSP3.  
 DR PIR; A53437; A53437.  
 DR MGD; MGI:97822; Pvr12.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_V.

RA Ling V., Dunnesi-Joannopoulos K.;  
 CC "G150 molecules and uses thereof."  
 RT Patent number WO0121796, 29-MAR-2001.  
 CC -I- FUNCTION: LIGAND FOR THE T-CELL-SPECIFIC CELL SURFACE RECEPTOR  
 CC ICOS. ACTS AS A COSTIMULATORY SIGNAL FOR T-CELL PROLIFERATION AND  
 CC CYTOKINE SECRETION; INDUCES ALSO B-CELL PROLIFERATION AND  
 CC DIFFERENTIATION INTO PLASMA CELLS. COULD PLAY AN IMPORTANT ROLE IN  
 CC MEDIATING LOCAL TISSUE RESPONSES TO INFLAMMATORY CONDITIONS, AS  
 CC WELL AS IN MODULATING THE SECONDARY IMMUNE RESPONSE BY CO-  
 CC STIMULATING MEMORY T CELL FUNCTION. DURING PREGNANCY, MAY FUNCTION  
 CC TO SKEW THE CYTOKINE OF MATERNAL T-CELLS TOWARD IMMUNOPROTECTIVE  
 CC TH2 PHENOTYPE.  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=1;  
 CC IsoId=Q9JH8-1; Sequence=Displayed;  
 CC Name=2; Synonyms=B;  
 CC IsoId=Q9JH8-2; Sequence=VSP\_002521;  
 CC -I- TISSUE SPECIFICITY: ISOFORM 1 HIGHEST EXPRESSION IN LYMPHOID  
 CC TISSUES, SUCH AS SPLEEN (MOSTLY IN THE MARGINAL ZONE), LYMPH NODES  
 CC (PARTICULARLY IN THE CORTX AND IN BOTH PRIMARY AND SECONDARY  
 CC FOLLICLES), THYMUS (PREDOMINANTLY IN THE MEDULLA) AND PETER'S  
 CC PATCHES (MOSTLY IN THE FOLLICLES); LOWER LEVELS IN MANY  
 CC NON-LYMPHOID TISSUES, SUCH AS BRAIN, HEART, KIDNEY, LIVER, LUNG,  
 CC SKELETAL MUSCLE AND TESTIS. PRESENT ON FRESHLY ISOLATED SPLENIC B-  
 CC CELLS, T-CELLS, DENDRITIC CELLS AND MACROPHAGES. THE EXPRESSION OF  
 CC ISOFORM 2 IS RESTRICTED TO HEART, SPLEEN AND KIDNEY.  
 CC -I- DEVELOPMENTAL STAGE: DETECTED EARLY IN EMBRYOGENESIS. IN THE YOLK  
 CC SAC AT 11.5 AND 12.5 DPC AND, TO A LESSER EXTENT, IN THE LIVER AT  
 CC 14.5 DPC.  
 CC -I- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. BTN/MOG  
 CC SUPERFAMILY.  
 CC -I- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -I- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF216747; AAF45149.1; -;  
 CC EMBL: AF199027; AAF34738.1; -;  
 CC EMBL: AX100591; CAC36463.1; -;  
 CC EMBL: AX100593; CAC36464.1; -;  
 CC EMBL: AF394451; AAK7544.1; -;  
 CC MGD: MGI:1354701; Icosl.  
 CC DR GO: GO:0005576; C:extracellular; IDA.  
 CC DR GO: GO:0016021; C:integral to membrane; NAS.  
 CC DR GO: GO:0003793; F:defense/immunity protein activity; NAS.  
 CC DR GO: GO:0005102; F:receptor binding activity; TAS.  
 CC DR GO: GO:0045190; P:isotype switching; NAS.  
 CC DR GO: GO:0045104; P:positive regulation of activated T-cell pro. .; TAS.  
 CC DR GO: GO:0045404; P:positive regulation of interleukin-4 biosyn. .; IMP.  
 CC DR GO: GO:0007165; P:signal transduction; NAS.  
 CC DR GO: GO:0042110; P:t-cell activation; NAS.  
 CC DR InterPro: IPR007110; IG-like.  
 CC DR InterPro: IPR003599; IG.  
 CC DR InterPro: IPR003006; IG\_MHC.  
 CC DR Pfam: PF00047; Ig\_1.  
 CC DR SMART: SM00409; Ig\_1.  
 CC DR PROSITE: PS50835; IG\_LIKE, 2.  
 CC B-cell activation; Immune response; Glycoprotein;  
 CC Immunoglobulin domain; Signal; Transmembrane; Multigene family;  
 CC Alternative splicing.  
 FT SIGNAL 1 46 BY SIMILARITY.  
 FT CHAIN 47 322 ICOS LIGAND.  
 FT DOMAIN 47 298 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 278 298 POTENTIAL.

FT DOMAIN 299 322 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 49 149 IG-LIKE V-TYPE.  
 FT DOMAIN 167 263 IG-LIKE C2-TYPE.  
 FT DOMAIN 31 38 POLY-LEU.  
 FT DOMAIN 289 292 POLY-ALA.  
 FT DISULFID 62 138 POTENTIAL.  
 FT DISULFID 185 243 POTENTIAL.  
 FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 130 120 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 213 213 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 321 322 HA -> TWAPVYQDYLLPRYLMSPLKTRGAP (in  
 FT isoform 2)  
 FT /FTID=VSP\_002521.  
 FT R -> H (IN REF. 4 AND 5; CAC36464).  
 FT CONFLICT 237 237  
 FT SEQUENCE 322 AA; 35960 MW; 55CCBA4AD12E4766 CRC64;  
 SQ  
 Query Match 7.4%; Score 118; DB 1; Length 322;  
 Best Local Similarity 24.0%; Pred. No. 0.0041;  
 Matches 63; Conservative 33; Mismatches 99; Indels 68; Gaps 11;  
 QY 11 LLLLLLLTTRSEVEYRAEYVGNAYLPCFYTPAPAGNL-VFVCGKGCAPVFCGNVYL 68  
 DB 35 LLLLSLCAASAEYEVGAMVGSNVVLSCLDPHRRHFNLSGLVYV-----QI 81  
 QY 69 RFDENDVNYW-----TSRYLNG-----DRKQDVSLTENVTLADSGICCRQ 113  
 DB 82 ENPEVSVTYTYLPEYKSGINVDSSYKRGHLSIDSKQGNFSLYLKNVLPQDQEFCTRV- 140  
 QY 114 IPGINDEKFNKLVIYKPAKTPAPTLQRFAPAPRLTTRGHGPAETQTL-----CSL 168  
 DB 141 ---FMTATELVKILEEYVRLVPAVF-----STPVISTSSNSPGGORTTCKSKCY 191  
 QY 169 PDINTQISTANELDRSLANDLSGATIRIGIYGAGICAGLALIFGALLPKWYS 228  
 DB 192 PEPNLYWINTDNSLIDTALQNT-----YLNK-----LGLVDVISTLLPMTS 236  
 QY 229 HSK-----EKT---QNLISLSLA 243  
 DB 237 RGDVLCVENVALHQNITSISQA 259  
 RESULT 5  
 A33\_HUMAN STANDARD; PRT; 319 AA.  
 ID A33\_HUMAN  
 AC Q99795;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Cell surface A33 antigen precursor (Glycoprotein A33).  
 GN GPA33.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RP TISSUE=Colon carcinoma;  
 RX MEDLINE=97165045; PubMed=9012807;  
 RA Heath J.K., White S.J., Johnstone C.N., Catimel B., Simpson R.J.,  
 RA Moritz R.L., Tu G.-F., U.H., Whitehead R.H., Groenen L.C.,  
 RA Scott A.W., Ritter G., Cohen L., Welt S., Old L.J., Nice B.C.,  
 RA Burgess A.W.;  
 RT "The human A33 antigen is a transmembrane glycoprotein and a novel  
 RT member of the immunoglobulin superfamily."  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:469-474(1997).  
 RN [2]  
 RP POST-TRANSLATIONAL MODIFICATIONS.  
 RX MEDLINE=97396159; PubMed=9245713;  
 RA Ritter G., Cohen L.S., Nice B.C., Catimel B., Burgess A.W.,



```
CC tyrosine kinase SYK.
CC - SUBUNIT: Interacts with TYROBP. This interaction results in the
CC recruitment of SYK.
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O00241-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O00241-2; Sequence=VSP_007026;
CC Note=No experimental confirmation available;
CC TISSUE SPECIFICITY: Detected in monocytes and dendritic cells.
CC - PTM: N-glycosylated.
CC - SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.
CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y10376; CAA71404.1; -
CC EMBL: AL049634; CAB4661.2; -
CC EMBL: AL138804; CAC1540.1; -
CC EMBL: BC025286; AAH25286.1; -
CC Genew: HGNC:15928; SIRPB1.
CC MIM: 603889; -.
CC GO: GO:0005887; C: integral to plasma membrane; TAS.
CC GO: GO:0007166; P: cell surface receptor linked signal transdu. .; TAS.
CC InterPro: IPR007110; Ig-like.
CC InterPro: IPR003597; Ig_C1.
CC InterPro: IPR003006; Ig_MHC.
CC Pfam: PF00047; Ig_3.
CC SMART: SM00407; IGc1.
CC PROSITE: PS50835; IG_LIKE; 3.
CC Repeat: Signal; Transmembrane; Immunoglobulin domain; Glycoprotein;
CC Alternative splicing.
CC FT SIGNAL 1 26 POTENTIAL.
CC FT CHAIN 27 398 SIGNAL-REGULATORY PROTEIN BETA-1.
CC FT DOMAIN 27 371 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 372 382 POTENTIAL.
CC FT DOMAIN 393 398 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 27 136 IG-LIKE V-TYPE.
CC FT DOMAIN 147 246 IG-LIKE C1-TYPE 1.
CC FT DOMAIN 253 347 IG-LIKE C1-TYPE 2.
CC FT DISULFID 54 120 POTENTIAL.
CC FT DISULFID 169 227 POTENTIAL.
CC FT DISULFID 272 320 POTENTIAL.
CC FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 269 269 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT VARSPLIC 145 361 Missing (in isoform 2).
CC FT CONFLICT 53 53 R -> H (IN REF. 2; CAC17540).
CC FT CONFLICT 102 102 D -> N (IN REF. 1).
CC FT CONFLICT 229 229 M -> I (IN REF. 2; CAB4661).
CC FT CONFLICT 363 363 P -> A (IN REF. 2).
CC FT CONFLICT 363 363 Missing (in isoform 2).
CC FT SEQUENCE 398 AA; 43255 MW; A2AA08FEB2B3C2B CRC64;
SO QUERY MATCH 7.5%; Score 118.5; DB 1; Length 398;
Best Local Similarity 24.2%; Pred. No. 0.0048;
Matches 75; Conservative 42; Mismatches 124; Indels 69; Gaps 16;
```

```
QY 106 GYICCRIDIPGIMANDEK-----NLKVIKP-AKVTPAFLQR--DTAAPPMLTTRG 156
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 116 GYIVCVKFRKSGPDDVEFKSGAGTELIVRAPSAVVGCPVRRAPRPHVTSP-----TCES 171
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 157 HGPARTQTLGSLPDINLQISTLANELDSRLANDLRSGATIRIGIYGICAGLALA 216
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 172 HG-----FSPRDTLTKWPKK-GNELSD--FQIVWDPAQDSVSYSHSTARV----- 214
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 217 LIFGALIFKWSHSEKTIQNLISLANIPSGGLANAAVEGISENITYTTEENYVEEE 276
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 215 -----VLIRGDVHS-QVICMAHITLQGDPLRGTMN-LSBAIRVPTL-EVYIQGPRAEN 266
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 277 PNEYICVYS 286
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 267 QANVTCQVSN 276
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
ICOL_MOUSE STANDARD; PRT; 322 AA.
AC Q9JHJ8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE ICOS ligand precursor (B7 homolog 2) (B7-H2) (B7-like protein G150)
DE (B7-related protein 1) (B7RP-1) (LICOS).
GN ICOSL OR B7H2 OR B7RP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RC TISSUE=Lymphocytes;
RX MEDLINE=20081495; PubMed=10617205;
RA Yoshinaga S.K., Whoriskey J.S., Khare S.D., Sarmiento U., Guo J.,
RA Horan T., Shin G., Zhang M., Coccia M.A., Kohno T., Tafuri-Bladt A.,
RA Brankow D., Campbell P., Chang D., Chiu L., Dai T., Duncan G.,
RA Elliott G.S., Hui A., McCabe S.M., Scully S., Shahinian A.,
RA Shaker C.L., Van G., Mak T.W., Senaldi G.;
RT "T-cell co-stimulation through B7RP-1 and ICOS.";
RL Nature 402:827-832(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Thymus;
RX MEDLINE=20015817; PubMed=10549624;
RA Swallow M.M., Mallin J.J., Sha W.C.;
RT "B7h, a novel costimulatory homolog of B7.1 and B7.2, is induced by
RT TNFalpha.";
RL Immunity 11:423-432(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C3H/HeJ; TISSUE=Fetal thymus;
RX MEDLINE=20126021; PubMed=10657606;
RA Ling V., Wu P.W., Finerney H.F., Bean K.M., Spaulding V., Fouser L.A.,
RA Leonard J.P., Hunter S.E., Zollner R., Thomas J.L., Miyashiro J.S.,
RA Jacobs K.A., Collins M.;
RT "Identification of G150, a novel B7-like protein that functionally
RT binds to ICOS receptor.";
RL J. Immunol. 164:1653-1657(2000).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Peripheral blood lymphocytes;
RX MEDLINE=21286479; PubMed=11390480;
RA Ling V., Wu P.W., Miyashiro J.S., Marusic S., Finerney H.F.,
RA Collins M.;
RT "Differential expression of inducible costimulator-1 ligand splice
RT variants: lymphoid regulation of mouse g150-b and human g150
RT molecules.";
RL J. Immunol. 166:7300-7308(2001).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
```

DR PDB; 1EAU; 13-JUL-01.  
DR PDB; 1FSW; 08-NOV-00.  
DR PDB; 1KAC; 24-NOV-99.  
DR Gene; HGNC:2559; CXADR.  
DR MIM; 602621; -.  
DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
DR GO; GO:0004872; F:receptor activity; TAS.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003598; IG\_C2.  
DR InterPro; IPR003006; IG\_MHC.  
DR Pfam; PF00047; Ig\_2.  
DR SMART; SM00408; IGC2\_1.  
DR PROSITE; PS00835; IG\_LIKE; 2.  
KW Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; signal;  
Repeat; 3D-structure.  
FT CHAIN 1 19 POTENTIAL.  
FT SIGNAL 20 365 COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR.  
FT TRANSMEM 238 258 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 259 365 POTENTIAL.  
FT DOMAIN 20 134 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 141 128 IG-LIKE C2-TYPE 1.  
FT DISULFID 41 120 IG-LIKE C2-TYPE 2.  
FT CARBOHYD 162 212 BY SIMILARITY.  
FT CARBOHYD 106 201 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 365 AA; 40029 MW; AB01C6346CB7FE64 CRC64;

Query Match 7.8%; Score 124.5; DB 1; Length 365;  
Beet Local Similarity 23.7%; Pred. No. 0.0012;  
Matches 61; Conservative 48; Mismatches 103; Indels 45; Gaps 12;

OY 9 CULLL-----LHLLTRSEVEYRAEYGVNAYLPCFT--PAAPGNLVPYCMGKACGP 59  
DB 6 CFWLLCGVDFPARSLSTITPEEMIEKAK-GERAYLPCKETLSPEQGLP-DIEMLISRPD 63  
OY 60 VPECNVLV--RTDERVNVYWTSR---YWLNGDFRKGDSVLTINVTLADSGIYCRR- 112  
DB 64 NQKQVQVILVSGDKYIDYDYPDLKGRVHFTSNDLKSQDSINVTNLQLSDIGTYQCKVK 123  
OY 113 QIRGIMNDEKFNLIKVIKPA---KYTPAPTLQIRP-----TAAFRMLTTRHGPAE 161  
DB 124 KAPGAN-KKIHVLVLPKPGARCYVDSEETGSGDFKIKCEPKESLPLQYEMQLSDSQ 182  
OY 162 TQTLGSLPDLNLTQIS-----TLANELRDSRLANDLRDQATIRIGIYGA 207  
DB 183 KMPSTWLMEMTSSVSVKASSESGYSCYVRANVSGDCLRLNVVPPSKAGLIGA 242  
OY 208 GICAGLALALIFGALIF 224  
DB 243 IIGTLALALILGLIF 258

RESULT 3  
SRBI\_HUMAN STANDARD; PRT; 398 AA.  
ID SRBI\_HUMAN O00241; O08T42; O9H1U5; O9Y4V0;  
AC 000241; O08T42; O9H1U5; O9Y4V0;  
DT 28-FEB-2003 (Rel. 41; Created)  
DT 15-SEP-2003 (Rel. 42; Last sequence update)  
DT 15-SEP-2003 (Rel. 42; Last annotation update)  
DE Signal-regulatory protein beta-1 precursor (SIRP-beta-1).  
GN SIRPB1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.  
OC NCI\_TaxID=9606;  
RX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Placenta;  
RX MEDLINE=97215901; PubMed=9062191;  
RA Kharitonov A., Chen Z., Sures I., Wang H., Schilling J.,  
RA Ullrich A.;  
RA "A family of proteins that inhibit signalling through tyrosine kinase

RT receptors.";  
RL Nature 386:181-186(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggaley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corry N.,  
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lehaeslalo M.H., Levesha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McCormachie L.J., McIay K., McMurtry A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Senra H.K., Showkhen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilkins L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemmer C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshlyuk S., Carinini P., Prange C.,  
RA Rahn S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosack S.A., McEwan P.J., McKernan K.J., Malek J.A., Gamaralle P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Wozny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M.,  
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smaltus D.E.,  
RA Scherich A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP INTERACTION WITH TYROBP AND SYK.  
RX PubMed=1169422;  
RA Tomassello E., Cant C., Buehring H.-U., Vely F., Andre P., Seifert M.,  
RA Ullrich A., Vavler E.;  
RT "Association of signal-regulatory proteins beta with KARAP/DAP-12.";  
RL Eur. J. Immunol. 30:2147-2156(2000).  
RN [5]  
RP INTERACTION WITH TYROBP, N-GLYCOSYLATION, AND TISSUE SPECIFICITY.  
RX MEDLINE=20072721; PubMed=10604985;  
RA Dietrich J., Cella M., Seifert M., Buehring H.-U., Colonna M.;  
RT "Signal-regulatory protein beta 1 is a DAP12-associated activating  
RT receptor expressed in myeloid cells.";  
RL J. Immunol. 164:9-12(2000).  
CC -!- FUNCTION: Immunoglobulin-like cell surface receptor involved in  
CC the negative regulation of receptor tyrosine kinase-coupled  
CC signaling processes. Participates also in the recruitment of

DR SMART; SMO0408; IGC2; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 2.  
 KW Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; signal;  
 Repeat.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 365 COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR  
 FT DOMAIN 20 237 HOMOLOG.  
 FT TRANSMEM 238 258 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 259 365 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 20 136 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 141 228 IG-LIKE C2-TYPE 2.  
 FT DISULFID 41 120 BY SIMILARITY.  
 FT DISULFID 162 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 201 365 VAAPNLSRMGAVPMVPMPSKSGSLV -> FYAVKTDGILT  
 FT CONFLICT 340 365 VV (IN REF. 2 AND 3).  
 SQ SEQUENCE 365 AA; 39947 MW; 5445B4B52A34B2A2 CRC64;  
 Query Match 8.3%; Score 131.5; DB 1; Length 365;  
 Best Local Similarity 23.3%; Pred. No. 0.00029;  
 Matches 65; Conservative 47; Mismatches 96; Indels 71; Gaps 14;  
 QY 9 CVLLLL-----LLLTSSSEVYRAVGNATLPCFTT--PAAPGLVAVCMKGAC 58  
 DB 6 CPLLGLGADFTSGISITTPGRIE-KAK-GETAVLPCKFTLSPDQGL-DIEW----- 57  
 QY 59 PVFECCGVVLTREDVDV-----NYMTSR-----YVLNGDPFKGVSLTIENV 100  
 DB 58 -----LISPDNQIVQVILLYSGDKYDNYIDPKRKHVFTSDVKGASINVTNL 110  
 QY 101 TLADSGIYCCRI-QIPGIMNDEKPNLKVIKPAK---VTPAPTLQRF-----TAAF 148  
 DB 111 QLSIDIGTQCKYKAPGVAN-KKFLTLTVKPSGRFCFVDGSEIEGNDPKLKCEPREGSL 169  
 QY 149 PPMLTTRHGPAETQTLGSLPINITLQIS-----TLANELRDRRLANDLRD 194  
 DB 170 PLQFQWQKLSQDTPTPWLAMTSPVLSVKNASSEYSGTYSCTVQNRVGSQCMRLDVL 229  
 QY 195 SGATIRIGIYIGAGICAGLALIFGALIFKMYSHSKK 233  
 DB 230 VPPSNRAGTIGAVIGTLLALVLI-GALIFCCHRRKREE 267  
 RESULT 2  
 CXXR\_HUMAN STANDARD; PRT; 365 AA.  
 ID P78310; O00694;  
 AC 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Coxsackievirus and adenovirus receptor (Coxsackievirus B-  
 adenovirus receptor) (hCAR) (CVB3 binding protein).  
 GN CXXAR OR CAR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97190109; PubMed=9036860;  
 RA Beigelson J.M., Cunningham J.A., Drogue G., Kurt-Jones E.,  
 RA Krithivas A., Hong J.S., Horwitz M.S., Crowell R.L., Flanberg R.W.;  
 RT "Isolation of a common receptor for Coxsackie B viruses and  
 RT adenoviruses 2 and 5.";  
 RT Science 275:1320-1323 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97250541; PubMed=9096397;  
 RA Tomko R.P., Xu R., Phillips L.;  
 RT "hCAR and hCAR: the human and mouse cellular receptors for subgroup C  
 RT adenoviruses and group B coxsackieviruses.";

RL Proc. Natl. Acad. Sci. U.S.A. 94:3352-3356 (1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20008750; PubMed=10543405;  
 RA Bowles K.R., Gibson J., Wu J., Shaffer L.G., Towbin J.A.,  
 RA Bowles N.E.;  
 RT "Genomic organization and chromosomal localization of the human  
 RT Coxsackievirus B-adenovirus receptor gene.";  
 RT Hum. Genet. 105:354-359 (1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX Anderson C.W., Kielczawa J., Dunn J., Freimuth P.;  
 RT "Sequence and expression of CXADR, the human gene for the  
 RT coxsackievirus and adenovirus receptor.";  
 RT Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX Anderson B., Tomko R., Andersson K., Darban H., Oncu D., Mizra M.,  
 RX Solterprant K., Sonhammer E., Phillips L.;  
 RT "Putative regulatory domains in the human and mouse CAR genes.";  
 RT Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -1- FUNCTION: SERVES AS A RECEPTOR FOR GROUP B COXSACKIEVIRUSES AND  
 CC SUBGROUP C OF ADENOVIRUSES (AD2 AND AD5).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Y07593; CAA6868.1; -;  
 DR EMBL; U90716; AAC51234.1; -;  
 DR EMBL; AF169366; AAF05908.1; -;  
 DR EMBL; AF169360; AAF05908.1; JOINED.  
 DR EMBL; AF169361; AAF05908.1; JOINED.  
 DR EMBL; AF169362; AAF05908.1; JOINED.  
 DR EMBL; AF169363; AAF05908.1; JOINED.  
 DR EMBL; AF169364; AAF05908.1; JOINED.  
 DR EMBL; AF169365; AAF05908.1; JOINED.  
 DR EMBL; AF200465; AAF24344.1; -;  
 DR EMBL; AF242865; AAG01088.1; -;  
 DR EMBL; AF242862; AAG01088.1; JOINED.  
 DR EMBL; AF242864; AAG01088.1; JOINED.  
 DR EMBL; BC003684; AAH03684.1; -;  
 DR EMBL; BC010536; AAH10536.1; -;

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 01:50:35 ; Search time 14.4811 Seconds  
(without alignments)  
977.464 Million cell updates/sec

Title: US-10-004-633-24

Perfect score: 1567  
Sequence: 1 MSHLPDCCVLLLLLLTR.....CVSSRQPSQPLGCRAMP 301

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	131.5	8.3	365	1	CGAR_MOUSE
2	124.5	7.8	365	1	CGAR_MOUSE
3	118.5	7.5	398	1	SRL1_HUMAN
4	118	7.4	322	1	ICOL_MOUSE
5	112.5	7.1	319	1	A33_HUMAN
6	109	6.9	530	1	PVR2_MOUSE
7	108	6.8	1138	1	TIR1_HUMAN
8	107	6.7	310	1	HEMA_RACV1
9	102.5	6.5	555	1	C166_CARAV
10	101.5	6.4	538	1	PVR2_HUMAN
11	100.5	6.3	387	1	SRL2_HUMAN
12	99	6.2	818	1	TRKB_CHICK
13	98.5	6.2	245	1	MOG_FAT
14	98	6.2	526	1	BUTY_HUMAN
15	97.5	6.1	315	1	HEMA_VACCT
16	97	6.1	506	1	SHS1_BOVIN
17	95	6.0	467	1	SIL5_MOUSE
18	95	6.0	503	1	SHS1_HUMAN
19	94.5	6.0	246	1	MOG_MOUSE
20	94.5	6.0	773	1	PIGR_RABIT
21	94	5.9	215	1	CIB2_RAT
22	94	5.9	246	1	MOG_BOVIN
23	93.5	5.9	349	1	CEA6_HUMAN
24	93	5.9	313	1	HEMA_VARV
25	92.5	5.8	306	1	CD60_MOUSE
26	92	5.8	1010	1	CONT_CHICK
27	92	5.8	1134	1	TIR1_MOUSE
28	91.5	5.8	459	1	CD4_RABIT
29	91	5.7	344	1	CEA6_HUMAN
30	90.5	5.7	519	1	ECTO_RAT
31	89.5	5.6	302	1	ICOL_HUMAN
32	89.5	5.6	314	1	HEMA_VACCV
33	89.5	5.6	344	1	NTRI_HUMAN

34	89	5.6	564	1	C166_BRARE	Q90460 brachydanio
35	88.5	5.6	344	1	NTRI_RAT	Q62718 ratu mus novy
36	88.5	5.6	771	1	PIGR_MOUSE	Q60939 homo sapien
37	88	5.5	215	1	CIB2_HUMAN	P20978 vaccinia vi
38	88	5.5	315	1	HEMA_VACCV	Q951h0 pan troglod
39	88	5.5	597	1	SIL1_PANTR	Q991j0 mus musculu
40	87.5	5.5	344	1	NTRI_MOUSE	Q63994 mus musculu
41	87.5	5.5	403	1	CD33_MOUSE	P13688 homo sapien
42	87.5	5.5	526	1	CEA1_HUMAN	P08714 vaccinia vi
43	86.5	5.5	315	1	HEMA_VACCV	Q61490 mus musculu
44	86.5	5.5	583	1	C166_MOUSE	Q60469 homo sapien
45	86	5.4	2012	1	DSOA_HUMAN	

## ALIGNMENTS

```
RESULT 1
CXAR_MOUSE          STANDARD;          PRT;          365 AA.
ID AC P97752; 009052;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Coxsackievirus and adenovirus receptor homolog precursor (mCAR).
GN CXADR OR CAR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=97190109; PubMed=9036860;
RA Bergelson J.M., Cunningham J.A., Droguett G., Kurt-Jones E.,
RA Krithivas A., Hong J.S., Horwitz M.S., Crowell R.L., Finberg R.W.;
RT "Isolation of a common receptor for Coxsackie B viruses and
RT adenoviruses 2 and 5."
RT Science 275:1320-1323 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/MA;
RX MEDLINE=97250541; PubMed=9096397;
RA Tomko R.P., Xu R., Philipson L.;
RT "HCAR and mCAR: the human and mouse cellular receptors for subgroup C
RT adenoviruses and group B coxsackieviruses."
RT Proc. Natl. Acad. Sci. U.S.A. 94:3352-3356 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RA Bergelson J.M., Krithivas A., Crowell T.L., Finberg R.W.;
RT "The murine CAR homologue (mCAR) is a receptor for coxsackie B
RT viruses and adenoviruses."
RT Submitted (May-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; Y10320; CAA71368.1; -
CC EMBL; U90715; AAC53148.1; -
CC EMBL; Y11929; CAA72679.1; -
CC MGD; MGI:1201679; Cxadr.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig; 2.
```

A: Molecule type: mRNA  
A: Residues: 1-315 <JIN>

A: Cross-references: GB:X15709; GB:M5773; NID:G61313; PID:CAA33740.1; PID:G61314

C: Superfamily: vaccinia virus hemagglutinin; immunoglobulin homology

C: Keywords: glycoprotein; hemagglutinin; late protein; transmembrane protein

F: 1-16/Domain: signal sequence #status predicted <SIG>

F: 17-315/Product: hemagglutinin #status predicted <MAT>

F: 27-105/Domain: immunoglobulin homology <IMM>

F: 280-302/Domain: transmembrane #status predicted <TM>

F: 303-315/Domain: intracellular #status predicted <INT>

F: 37,69,112,161,254/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.1%; Score 97.5; DB 1; Length 315;

Best Local Similarity 17.4%; Pred. No. 1.1; Indels 97; Gaps 12;

Matches 57; Conservative 57; Mismatches 116; Indels 97; Gaps 12;

QY 3 SHLPDCVLLLLLTSSSEVEYRAEVGQNAVLPCEFTPAAPGN-----LVPVCMGK 55

DB 2 ARLP---ILLILSLVSTPSPQTSKIDDLTSC-----NRNNDYVVMASMYK 50

QY 56 GACPVFECGNVLTDERDV---NYWTSRYWANGDFRKGVSLTIENVTLADSGIYCCR 111

DB 51 -----EPNSITLLAKSDVLFEDNYTKKISYDSPYDDLVTITIKSLTARDAGTYCA 104

QY 112 IOIPGIMD-----EKFKLVIVIPAKVTPAFTLORDFTAAPFRLTLTGHPAETQL 165

DB 105 FMTSPPTDTRKVDYEEYESTELIV-----NTDESTIDIILSGSTHSP---ETS 150

QY 166 GSLP-DINLTQISTLANELRDSRLANDLDSGATIRIGIYGAGICAGLALIFGALIF 224

DB 151 SEKPEDIDVNCSSVFEITPPTIDNVEDHTDTV----- 185

QY 225 KWSHSEKIQNLISLISLANLPPSGLANAVAGIRSEENITYIENV----- 271

DB 186 ---TYSDSINTVSATS-----GESTDETPETIDKEEDHTVPTVSSTVSTSGIYT 237

QY 272 --YVEEPNEYCYVSSRQPSQPLGC 296

DB 238 TKSTDDADLYDPTVNDNTVPTTVGC 264

#### RESULT 15

T28598

hypothetical protein J7R - variola major virus

C: Species: variola major virus

C: Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #ext\_change 21-Jul-2000

C: Accession: T28598

R: Masung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin

A: Title: Potential virulence determinants in terminal regions of variola smallpox virus

A: Reference number: Z20488; PMID:94088747; PMID:8264798

A: Accession: T28598

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Cross-references: EMBL:L22579; NID:G623595; PID:AAA60908.1; PID:G439077

A: Experimental source: strain Bangladesh 1975

C: Superfamily: vaccinia virus hemagglutinin; immunoglobulin homology

Query Match 6.0%; Score 96; DB 2; Length 313;

Best Local Similarity 21.5%; Pred. No. 1.4; Indels 42; Gaps 8;

Matches 38; Conservative 35; Mismatches 62; Indels 42; Gaps 8;

QY 11 LLLLLLTSS---EVEYRAEVGQNAVLPCEFTPAAPGNL---VPCMGKACPFEC 63

DB 6 ILLILSLVSTPPIQISKIGDDATLSC-----SRNNNDYVVMASMYK-----EP 54

QY 64 GNVVLTDERDV---NYWTSRYWANGDFRKGVSLTIENVTLADSGIYCCRITQIPGIM 119

DB 55 NSIILAAKSDVLYFEDNYTKKISYDSPYDDLVTITIKSLTARDAGTYVCAEFMTSTN 114

QY 120 D-----EKFKLVIVIPAKVTPAFTLORDFTAAPFRLTLTGHPAETQLGLSPD 170

DB 115 DTDKVDYEEYESTELIV-----NTDESTIDIILSGSSHSP---ETSSEKPD 157

Search completed: November 22, 2003, 05:44:26

Job time: 27.8247 secs

A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-818 <V12>  
A:Cross-references: EMBL:X77251; NID:9563881; PIDN:CAA54468.1; PID:g472934  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1994  
R:Dechant, G.; Biffo, S.; Okazawa, H.; Kolbeck, R.; Potgiesser, J.; Barde, Y.A.  
A>Title: Expression and binding characteristics of the BDNF receptor chick trkB.  
A:Reference number: S42175; MUID:94116452; PMID:8287802  
A:Accession: S42175  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-818 <DEC>  
A:Cross-references: EMBL:X74109; NID:g407798; PIDN:CAA52210.1; PID:g407799  
A:Gene: trkB  
A:Function:  
A:Description: regulation of nervous system development; receptor for brain-derived neurotrophic factor  
C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glycoprotein  
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor kinase  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-818/Product: brain-derived neurotrophic factor receptor #status predicted <MAT>  
F:32-432/Domain: extracellular #status predicted <EXT>  
F:66-90/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:91-114/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:115-137/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:138-149/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical <LR>  
F:433-449/Domain: transmembrane #status predicted <TM>  
F:450-818/Domain: cytosolic #status predicted <CT>  
F:532-810/Domain: protein kinase homology <KIN>  
F:540-548/Region: protein kinase ATP-binding motif  
F:566,94,120,199,204,253,287,324,337,349,408/Binding site: carbohydrate (Asn) (covalent)  
F:566/Active site: lys #status predicted  
F:702/Binding site: phosphate (Tyr) (covalent) #status predicted  
F:813/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 6.2%; Score 99; DB 1; Length 818;  
Best Local Similarity 19.9%; Pred. No. 2.6;  
Matches 59; Conservative 41; Mismatches 89; Indels 108; Gaps 12;

QY 87 DPRKGDVSLITENVTLASGIY--CCRIQIPG-INNDEKF-----NLKLVTKP----- 131  
DB 241 DTSKNPASTLTKNVSMSDGLWISCAVENIVGEVOTSALTFVFAPIITFIESPDPDHW 300

QY 132 -----AKTTPATLDRDFTFAAPR-----M 151  
DB 301 CIPFTVKGKPKFTLQWTFBGLINSEYICTKIHIINOSEYHGCLOLDNPHLNGAYTL 360

QY 152 LTRGHGPAETQT-----LGSLPDINI/IOISTLANELRDSR-----LAN 190  
DB 361 LAKNEYGEDEKRVADHFMSPVPGDSGPVDPVVEYETTPNDLGDITNNSNGITSPDVSN 420

QY 191 DLKDSGATIRIGIYAGAGICAGLALALIFGALIFPMYSHSKKIKONTLSIT----- 240  
DB 421 KENEDSTIVYVVGVAIAVCTGLVIMLI-----ILKFRGRSHKFGMGPPSSVLSNDDSDASP 476

QY 241 -----SLANLPPS--GLANAVAEGIRSEENIYTIENYVEPEPEHYCVSSROOP 290  
DB 477 LHHISNGSTPSSSEGGPDVAITGM-----TKIFVIEHP-QYFQITNSQLKP 522

RESULT 12  
161783  
sodium channel beta 2 subunit - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: I61783  
R:Isom, L.L.; Ragsdale, D.S.; DeGough, K.S.; Westendbroek, R.E.; Reber, B.F.X.; Scheuer, A.; Title: Structure and Function of the Beta 2 Subunit of Brain Sodium Channels, a Transm  
A:Reference number: AS7843; MUID:96067641; PMID:8521473  
A:Accession: I61783

A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-186 <RES>  
A:Cross-references: EMBL:U37147; NID:g1086498; PIDN:AA60506.1; PID:g1086499  
A:Gene: SCNB2  
A:Introns: 50/3; 120/2

Query Match 6.2%; Score 98; DB 2; Length 186;  
Best Local Similarity 22.9%; Pred. No. 0.5;  
Matches 50; Conservative 21; Mismatches 69; Indels 78; Gaps 9;

QY 31 GQNAVLPCTYPAAPGN--LVPVCGWKGACPVFECGN-----VLNTRDRDVNWTSSRYW- 83  
DB 14 GSDTRLPCTFMNSCYVNHKQFSIWN-----TYQESNCSSEENVLQFRKIKIINLKERGD 68

QY 84 ---LNGDRKGDVSLITENVTLASGIYCCRIQIPGINNDEKFNKLVYKPAKTPAPTL 140  
DB 69 RVFESGNSKVDVSTLKNVQLEDEGINCYITNP-----PDR----- 106

QY 141 QRDFAAPRMLTTGHPAETQTGSLPDINL/IOISTLANELRDSRLANDLSDGATIR 200  
DB 107 -----HRGSKITVQLVLEVP-----PERDSTVAIT 132

QY 201 IGIVYAGICAGLALALIFGALIFPMYSHSKKIKONTLS 238  
DB 133 VGASVG-----GFLAVVILVLMVVCVARKE--QKLS 163

RESULT 13  
570587  
butyrophilin precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 21-Jul-2000  
C:Accession: S70587  
R:Taylor, M.R.; Peterson, J.A.; Ceriani, R.L.; Couto, J.R.  
Biochim. Biophys. Acta 1306, 1-4, 1996  
A>Title: Cloning and sequence analysis of human butyrophilin reveals a potential recepto  
A:Reference number: S70587; MUID:96201696; PMID:8611614  
A:Accession: S70587  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-526 <TAY>  
A:Cross-references: EMBL:U39576; NID:g1326082; PIDN:AAC50489.1; PID:g1326083

Query Match 6.2%; Score 98; DB 2; Length 526;  
Best Local Similarity 28.4%; Pred. No. 1.8;  
Matches 40; Conservative 21; Mismatches 58; Indels 22; Gaps 6;

QY 3 SHLPDVCYLILLILLITRSSSEY-----RAVGNQNAVLPCTYPAAPGNLVPVCG 53  
DB 7 SGLP-RCLTLTLLIQPLKDSAPFPVGPPEPIIYAVGEDDELPERLSPNMSAEHLRW 65

QY 54 -GKACPVFECGNVLTDRD-----VNVTSSRYWNGD-FRKGDVSLITENVTLASDG 106  
DB 66 FRKQVSPA-----VLVHRDGRQEAQEPYRGATLVQDGIAGKRVALLRIGVAVSDG 120

QY 107 IYCCRIQIPGINNDEKFNKLV 127  
DB 121 EYTCFREDGSYEALVHLKV 141

RESULT 14  
HNVZVT  
hemagglutinin precursor - vaccinia virus (strain Tianan)  
C:Species: Vaccinia virus  
C:Date: 30-Jun-1990 #sequence\_revision 30-Jun-1990 #text\_change 16-Jul-1999  
C:Accession: J0108  
R:Jin, D.; Li, Z.; Jin, O.; Yuwen, H.; Hou, Y.  
J.Exp. Med. 170, 571-576, 1989  
A>Title: Vaccinia virus hemagglutinin. A novel member of the immunoglobulin superfamily.  
A:Reference number: J0108; MUID:89328331; PMID:2754392  
A:Accession: J0108

```

Db      384 KRRSSDGKHKVLTCAEBSGPKPDVQSVNGINDVSYNNKA---TYKLTIVPSKNTLVS 440
      177 STLANEL-----RDSRLANDLRDGS---ATIRIGIYGAGICAGLALIFGALIFR----- 225
Qy      441 CLVTIKLGBDTRKEISVFSGKNEDGTEQAKVIVIGVLLVAAL-VGIIYWIYIKKTGG 499
Db      226 -WYSHSK 232
      500 SWKTGEKE 507

```

```

RESULT 8
153960
PRR2 alpha - human
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C:Accession: 153960
R:Eberle, F.; Dubreuil, P.; Mattei, M.G.; Devillard, E.; Lopez, M.
Gene 159, 267-272, 1995
A>Title: The human PRR2 gene, related to the human poliovirus receptor gene (PVR), is ch
A:Reference number: 153960; MUID:95347610; PMID:7622062
A:Accession: 153960
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-478 <RES>
A:Cross-references: GB:S79171; NID:G1042202; PID:G1042203
C:Genetics:
A:Gene: PRR2alpha
C:Superfamily: poliovirus receptor; immunoglobulin homology
F:276-331/Domain: immunoglobulin homology <IMM>

```

```

Query Match      6.4%; Score 101.5; DB 2; Length 478;
Best Local Similarity 25.6%; Pred. No. 0.81;
Matches 57; Conservative 29; Mismatches 96; Indels 41; Gaps 13;

Qy      11 LLLLLLLTRSEV-----EYRAVGQNAVLPCEYTPAAGNLVP-VCMGKACPVFEC 63
      21 LLLLLLLETGADVRVQVLPVVRGQLGGTVLPCHLPLPVGLVLSLVTWQRDPAPA-NH 79
Db      64 GNVVL-----RTDERDVNYWTSRYWLNDFRK--GDVSLTIENVTLADSGIYC 109
      80 QNVAAFHFKMGSPSPKSGSERLSFVSAKOSTGQDTAELODATALHGLTVEDEGNYT 139
Qy      110 CRI-----QIPGINDEKFNKLVIRPAKVPAP--TLQDFTAAPFPMULTTRGHGPAE 161
      140 CEFATFPKGSVRGWT-----WLRVIAPKPKQAQAQKVTFSODPTTV-ALCISKEGPPAR 193
Db      162 TOTLGSLLP-DINLTQIS-TLANELR-DSR--LANDLRDGGATI 199
      194 ISWLSLMEAKETQVSGTLAGTVVTSRFTLVPSGRADGTV 236

```

```

RESULT 9
168093
PRR2 delta - human
C:Species: Homo sapiens (man)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jan-2000
C:Accession: 168093
R:Eberle, F.; Dubreuil, P.; Mattei, M.G.; Devillard, E.; Lopez, M.
Gene 159, 267-272, 1995
A>Title: The human PRR2 gene, related to the human poliovirus receptor gene (PVR), is ch
A:Reference number: 153960; MUID:95347610; PMID:7622062
A:Accession: 168093
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-538 <RES>
A:Cross-references: GB:S79172; NID:G1042204; PID:G1042205
C:Genetics:
A:Gene: PRR2delta
C:Superfamily: poliovirus receptor; immunoglobulin homology
F:276-331/Domain: immunoglobulin homology <IMM>

```

```

Query Match      6.4%; Score 101.5; DB 2; Length 538;
Best Local Similarity 25.6%; Pred. No. 0.93;
Matches 57; Conservative 29; Mismatches 96; Indels 41; Gaps 13;

Qy      11 LLLLLLLTRSEV-----EYRAVGQNAVLPCEYTPAAGNLVP-VCMGKACPVFEC 63
      21 LLLLLLLETGADVRVQVLPVVRGQLGGTVLPCHLPLPVGLVLSLVTWQRDPAPA-NH 79
Db      64 GNVVL-----RTDERDVNYWTSRYWLNDFRK--GDVSLTIENVTLADSGIYC 109
      80 QNVAAFHFKMGSPSPKSGSERLSFVSAKOSTGQDTAELODATALHGLTVEDEGNYT 139
Qy      110 CRI-----QIPGINDEKFNKLVIRPAKVPAP--TLQDFTAAPFPMULTTRGHGPAE 161
      140 CEFATFPKGSVRGWT-----WLRVIAPKPKQAQAQKVTFSODPTTV-ALCISKEGPPAR 193
Db      162 TOTLGSLLP-DINLTQIS-TLANELR-DSR--LANDLRDGGATI 199
      194 ISWLSLMEAKETQVSGTLAGTVVTSRFTLVPSGRADGTV 236

```

```

RESULT 10
JC2457
vascular cell adhesion protein - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 15-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 08-Oct-1999
C:Accession: JC2457
R:Tsang, Y.T.M.; Haskard, D.O.; Robinson, M.K.
Biochem. Biophys. Res. Commun. 201, 805-812, 1994
A>Title: Cloning and expression kinetics of porcine vascular cell adhesion molecule.
A:Reference number: JC2457; MUID:94271236; PMID:7516159
A:Accession: JC2457
A:Molecule type: mRNA
A:Residues: 1-538 <TSA>
A:Cross-references: EMBL:U08351; NID:G474382; PIDN:AAA21542.1; PID:G474383
C:Keywords: glycoprotein; transmembrane protein
F:497-517/Domain: transmembrane #stratus predicted <TM>
F:75,157,271,330,360/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

```

Query Match      6.4%; Score 101; DB 2; Length 538;
Best Local Similarity 25.6%; Pred. No. 1;
Matches 61; Conservative 29; Mismatches 78; Indels 70; Gaps 13;

Qy      85 NGD--FRKGDVSLTIENVTLADSGIYCCRIQIPGINDEKFNKLVIRPAKVPAPTLOR 142
      262 NGDQQLLSGNATLTLLIARMEDSGIYCEGVNPVGTNRKEVELTVQVAP-----R 311
Db      143 DETAAPFPMULTTRGHGPAETQTLGSLPDINTQIS-----TLANELRDSRL----- 188
      312 D-----TTISVNPSTLEBGS--SVNMTCSGSDGFPAPKILMSKLRDGNLPLESEN 360
Qy      189 -----ANDLRDGGATIRIGIYGAGI-CAGL--ALALIFGA-----LIFKWSHSEKI 234
      361 TTLTLTSTRKEDS-----GIYCEGINQAQINRKEVELLIQAAPKDLQTLTFPESVYE 414
Db      235 QNLSLIS--LANLPSGL-----ANAAVAGIRSENIYTI-----EENVYVEEPPNE 279
      415 GDTVILISCTCGVPEPTLLILKKKAKETGDTVLKSTGAVTIHRARLADGAVYECESKNE 472

```

```

RESULT 11
S44098
brain-derived neurotrophic factor receptor precursor - chicken
N:Alternate names: receptor tyrosine kinase trkB
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Gallus gallus (chicken)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-Jun-1999
C:Accession: S59939; S42175; S44098
R:Vinik, N.O.; Erdmann, K.S.; Heumann, R.
Gene 149, 383-384, 1994
A>Title: Cloning and sequence analysis of a cDNA encoding a novel truncated form of the
A:Reference number: S59939; MUID:95047511; PMID:7959025
A:Accession: S59939

```





```
QY      247  P S G L A N A V A E G R S E N I Y T T E E N V E E E      276
          | | | | : : : | | | : : :
Db      412  F K T L Q N A V K K E V H A E D N I Y - I E N N L A M N Q      440
```

## RESULT 2

UC7780  
 C/Accession: JC7780  
 C/Species: Bos primigenius taurus (cattle)  
 C/Date: 02-Apr-2002 #sequence\_revision 02-Apr-2002 #text\_change 02-Apr-2002  
 R/Tinoclen, I.; Keyaerts, E.; Lindberg, M.; Van Ranst, M.  
 Biochem. Biophys. Res. Commun. 288, 805-808, 2001  
 A/Title: Characterization of a cDNA encoding the bovine coxsackie and adenovirus receptor  
 A/Reference number: JC7780  
 A/Contents: Liver  
 A/Accession: JC7780  
 A/Molecule type: mRNA  
 A/Residues: 1-365 <THO>  
 A/Cross-references: GB:AY033651  
 A/Comment: This protein serves as the primary adenoviral attachment site on bovine cells

Quest Match	7.6%	Score	120;	DB	2;	Length	365;
Quest Local	23.9%	Pred.	No. 0.015;				
Matches	64;	Conservative	44;	Mismatches	110;	Indels	50;
						Gaps	13

```

Qy 11 LLLLLLLL-----TSSSEVEYRAEV-----GNNYLCEPFT--PAAAGNIVPVMCKGA 57
Db 3 LLLRFLLLCGVADFTRGLSITTPQGMIEKAKETAYLVCKRTLGPEDGGPL-DIEWLSP 61

Qy 58 CPVPEGCGNYVL-RTDERDVNTWTS---YWLNGDFPKGDVSLTIEVNTLADSGIYCCR 111
Db 62 ADNQAKDQVDFIILYSGKIDYDYOCLKGRVHFTSDMLSGDASINVNMLQLSDIGTYQCK 121

Qy 112 I-OIPGIMDEKENLGLVAKPA---KYTPAPTLQDF-----TAAFPMLTTRGHG 159
Db 122 VKRAPGVGN-KIQLVLVVKPSGIRCYNDGSEIENDPKLCEPREGSLPIRYWQQLSD 180

Qy 160 AETQTLGSLPDINLTOIS-----TLANELDRSLRLANDLSDGATIRIGIYI 205
Db 181 SQKLPTSMPLPEMTSPVIVKMASEAEGTYGTCVAVNRVGSQCLRLDVPVPSNRAGTIA 240

Qy 206 GAGICAGLALALIFGLALFKWYSHSEK 233
Db 241 GAVI--GTLALVLALIV-FCCHKRR 265

```

### RESULT 3

poliovirus receptor homolog precursor - mouse  
C:Species: Mus musculus domesticus (western European house mouse)  
C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 22-Jun-1999  
C:Accession: A38211  
R:Morrisson, M.E.; Racaniello, V.R.  
J. Virol. 66, 2807-2813, 1992  
A:Title: Molecular cloning and expression of a murine homolog of the human poliovirus re  
A:Reference number: A38211; MUID:92219365; PMID:1560525  
A:Accession: A38211  
A:Molecule type: DNA  
A:Residues: 1-467 <MOR>  
A:Cross-references: GB:M60206; NID:G199785; PID:AAA39734.1; PID:G199786  
C:Superfamily: poliovirus receptor; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; transmembrane protein  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-467/Product: poliovirus receptor homolog #status predicted <MAT>  
F:26-354/Domain: extracellular #status predicted <EXT>  
F:47-133/Domain: immunoglobulin homology <IMM1>  
F:167-231/Domain: immunoglobulin homology <IMM2>  
F:267-322/Domain: immunoglobulin homology <IMM3>  
F:355-374/Domain: transmembrane #status predicted <TMN>  
F:375-467/Domain: intracellular #status predicted <INT>  
E:54-131,174-229,274-320/Distal bonds: #status predicted  
F:129,138,315/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match	6.9%;	Score 109;	DB 1;	length 467;
Best Local Similarity	22.0%;	Pred. No. 0.18;		
Matches	63;	Conservative	33;	Mismatches 103;
			Indels	88;
			Gaps	10;

```

Qy 11 LLLLLLLLRSSEV-----EYRAVCGNALLPEFYTPAAGNVAIPCMCK-----GA 57
    |||||
Db 21 LLLLLLQETGADVARVRLPEVRGRLGTVELPCHLLPPTTERSQYTMGRDGTVAAF 80
Qy 58 CPPEFCGNVLTDERDNNYTSRWMLNGDRKXQDSLTIENTLADSGYICRI-QIPG 116
    |||||
Db 81 HPSFGVDPEPNSQFSXDRLSFVARARETADLR--DALTAFLGRGLVEEGNYTCEPATFPN 138
Qy 117 IMDEKENFKLTKP-----AKVTPAPTLQORDFLAAPPRLTLTRGHGPAETQTLGSLPD 170
    |||||
Db 139 GTRRGYTWLRVTAQGEENNAEAOEYTIQO-----SVAARCVSTGCGPPA----- 183
Qy 171 INTQISTLANLRDSRLANDLRDSGATIRIGIYIGACIGAGLALALFGALLFKWYSHS 230
    |||||
Db 184 -RTWISSISGGEAKDQO-----EPGIQAG-----206
Qy 231 KERIQWLSLISLANLPQGLANAVAEGRSENIYTTIEENVYEVEEP 277
    |||||
Db 207 -----TWTIISKRLSPVGRADGVAVYTRVHSEF-----EEP 239

```

## RESULT 4

poliovirus receptor mpvr - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Jun-2000  
 C:Accession: A53437  
 R:Acct1, J.; Koike, S.; Ise, I.; Sato-Yoshida, Y.; Nomoto, A.  
 J. Biol. Chem. 269, 8431-8438, 1994  
 A:Title: Amino acid residues on human poliovirus receptor involved in interaction with F  
 A:Reference number: A53437, PMID:8132559  
 A:Accession: A53437  
 A:Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-530 <AOK>  
 A:Cross-references: GB:J26107; MID:g475017; PIDN:BA05103.1; PID:g825507  
 A:Experimental source: C57/BL6, brain  
 A:Note: sequence extracted from NCBI backbone (NCBI:146664, NCBI:P.146667)  
 C:Superfamily: poliovirus receptor; immunoglobulin homology  
 I:47-135/Domain: immunoglobulin homology <IMM>

Query Match	6.9%;	Score 109;	DB 2;	Length 530;
Best Local Similarity	22.0%;	Pred. No. 0.21;		
Matches	63;	Conservative	33;	Mismatches 103;
				Indels 88;
				Gaps 10

```

QY      11  LLLLLLLLRSSEV-----EYRAVGVGNAAVPCGYTPAAGNVLPVCMGK-----GA  57
      12  |||||
      13  |||||
      14  |||||
      15  |||||
      16  |||||
      17  |||||
      18  |||||
      19  |||||
      20  |||||
      21  LLLLLLQEGADQVAVRVLPVYRGRGLGGTVELLPCHLBPPTTERRSQVTLWGRLDSTVVAAF  80
      22  |||||
      23  |||||
      24  |||||
      25  |||||
      26  |||||
      27  |||||
      28  |||||
      29  |||||
      30  |||||
      31  |||||
      32  |||||
      33  |||||
      34  |||||
      35  |||||
      36  |||||
      37  |||||
      38  |||||
      39  |||||
      40  |||||
      41  |||||
      42  |||||
      43  |||||
      44  |||||
      45  |||||
      46  |||||
      47  |||||
      48  |||||
      49  |||||
      50  |||||
      51  |||||
      52  |||||
      53  |||||
      54  |||||
      55  |||||
      56  |||||
      57  |||||
      58  CPVFECCGNVLRTERDEVNWTYSRYLNGDFRKGDVSLTIENVTLADSGIYCCRI-QIPG  116
      59  |||||
      60  |||||
      61  HPSFGVDVPNSQSKORLSFVARARPETNADLR--DATLAFRGLAVEDEGNVTCFATFPN  138
      62  |||||
      63  |||||
      64  |||||
      65  |||||
      66  |||||
      67  |||||
      68  |||||
      69  |||||
      70  |||||
      71  |||||
      72  |||||
      73  |||||
      74  |||||
      75  |||||
      76  |||||
      77  |||||
      78  |||||
      79  |||||
      80  |||||
      81  |||||
      82  |||||
      83  |||||
      84  |||||
      85  |||||
      86  |||||
      87  |||||
      88  |||||
      89  |||||
      90  |||||
      91  |||||
      92  |||||
      93  |||||
      94  |||||
      95  |||||
      96  |||||
      97  |||||
      98  |||||
      99  |||||
      100  |||||
      101  |||||
      102  |||||
      103  |||||
      104  |||||
      105  |||||
      106  |||||
      107  |||||
      108  |||||
      109  |||||
      110  |||||
      111  |||||
      112  |||||
      113  |||||
      114  |||||
      115  |||||
      116  |||||
      117  INNDEKFNKLVIKP-----AKVTPAPTLQRODLPAAFPKMLTTRGHGPAETQTLGSLPD  170
      118  |||||
      119  |||||
      120  |||||
      121  |||||
      122  |||||
      123  |||||
      124  |||||
      125  |||||
      126  |||||
      127  |||||
      128  |||||
      129  GTRGRGTWLRVIAQEPENHAEAOEVTIGPQ-----SVAARCVSGTCGRPPA-----  183
      130  |||||
      131  |||||
      132  |||||
      133  |||||
      134  |||||
      135  |||||
      136  |||||
      137  |||||
      138  |||||
      139  |||||
      140  |||||
      141  |||||
      142  |||||
      143  |||||
      144  |||||
      145  |||||
      146  |||||
      147  |||||
      148  |||||
      149  |||||
      150  |||||
      151  |||||
      152  |||||
      153  |||||
      154  |||||
      155  |||||
      156  |||||
      157  |||||
      158  |||||
      159  |||||
      160  |||||
      161  |||||
      162  |||||
      163  |||||
      164  |||||
      165  |||||
      166  |||||
      167  |||||
      168  |||||
      169  |||||
      170  |||||
      171  INTQSTLNLRLPDSRLANDLRDSGATIRIGYIGGICGLMALILFGALIFKWSHS  2319
      172  |||||
      173  |||||
      174  |||||
      175  |||||
      176  |||||
      177  |||||
      178  |||||
      179  |||||
      180  |||||
      181  |||||
      182  |||||
      183  |||||
      184  -RTTWISSLGGGEAKTQ-----EPGIQG-----  206
      185  |||||
      186  |||||
      187  |||||
      188  |||||
      189  |||||
      190  |||||
      191  |||||
      192  |||||
      193  |||||
      194  |||||
      195  |||||
      196  |||||
      197  |||||
      198  |||||
      199  |||||
      200  |||||
      201  |||||
      202  |||||
      203  |||||
      204  |||||
      205  |||||
      206  |||||
      207  |||||
      208  |||||
      209  |||||
      210  |||||
      211  |||||
      212  |||||
      213  |||||
      214  |||||
      215  |||||
      216  |||||
      217  |||||
      218  |||||
      219  |||||
      220  |||||
      221  |||||
      222  |||||
      223  |||||
      224  |||||
      225  |||||
      226  |||||
      227  |||||
      228  |||||
      229  |||||
      230  |||||
      231  |||||
      232  |||||
      233  |||||
      234  |||||
      235  |||||
      236  |||||
      237  |||||
      238  |||||
      239  |||||
      240  |||||
      241  |||||
      242  |||||
      243  |||||
      244  |||||
      245  |||||
      246  |||||
      247  |||||
      248  |||||
      249  |||||
      250  |||||
      251  |||||
      252  |||||
      253  |||||
      254  |||||
      255  |||||
      256  |||||
      257  |||||
      258  |||||
      259  |||||
      260  |||||
      261  |||||
      262  |||||
      263  |||||
      264  |||||
      265  |||||
      266  |||||
      267  |||||
      268  |||||
      269  |||||
      270  |||||
      271  |||||
      272  |||||
      273  |||||
      274  |||||
      275  |||||
      276  |||||
      277  |||||
      278  |||||
      279  |||||
      280  |||||
      281  |||||
      282  |||||
      283  |||||
      284  |||||
      285  |||||
      286  |||||
      287  |||||
      288  |||||
      289  |||||
      290  |||||
      291  |||||
      292  |||||
      293  |||||
      294  |||||
      295  |||||
      296  |||||
      297  |||||
      298  |||||
      299  |||||
      300  |||||
      301  |||||
      302  |||||
      303  |||||
      304  |||||
      305  |||||
      306  |||||
      307  |||||
      308  |||||
      309  |||||
      310  |||||
      311  |||||
      312  |||||
      313  |||||
      314  |||||
      315  |||||
      316  |||||
      317  |||||
      318  |||||
      319  |||||
      320  |||||
      321  |||||
      322  |||||
      323  |||||
      324  |||||
      325  |||||
      326  |||||
      327  |||||
      328  |||||
      329  |||||
      330  |||||
      331  |||||
      332  |||||
      333  |||||
      334  |||||
      335  |||||
      336  |||||
      337  |||||
      338  |||||
      339  |||||
      340  |||||
      341  |||||
      342  |||||
      343  |||||
      344  |||||
      345  |||||
      346  |||||
      347  |||||
      348  |||||
      349  |||||
      350  |||||
      351  |||||
      352  |||||
      353  |||||
      354  |||||
      355  |||||
      356  |||||
      357  |||||
      358  |||||
      359  |||||
      360  |||||
      361  |||||
      362  |||||
      363  |||||
      364  |||||
      365  |||||
      366  |||||
      367  |||||
      368  |||||
      369  |||||
      370  |||||
      371  |||||
      372  |||||
      373  |||||
      374  |||||
      375  |||||
      376  |||||
      377  |||||
      378  |||||
      379  |||||
      380  |||||
      381  |||||
      382  |||||
      383  |||||
      384  |||||
      385  |||||
      386  |||||
      387  |||||
      388  |||||
      389  |||||
      390  |||||
      391  |||||
      392  |||||
      393  |||||
      394  |||||
      395  |||||
      396  |||||
      397  |||||
      398  |||||
      399  |||||
      400  |||||
      401  |||||
      402  |||||
      403  |||||
```

## 4

RESULT 5  
S24066

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 05:17:26 / Search time 24.8247 Seconds  
(without alignments)  
1166.045 Million cell updates/sec

Title: US-10-004-633-24

Perfect score: 1587

Sequence: 1 MFSHLPFDCCVLLILLILLITR.....CVSSRQPSQPIGCRFAMP 301

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	256	16.1	451	2 S71754	cellular hepatitis
2	120	7.6	365	2 JC7780	coxsackie- and ade
3	109	6.9	467	1 HLMSP2	poliovirus recepto
4	109	6.9	530	2 A53437	poliovirus recepto
5	108	6.8	1138	1 S24066	protein-tyrosine k
6	107	6.7	310	1 HNVZRA	hemagglutinin prec
7	102.5	6.5	523	2 IS0478	neurofilin - goldfis
8	101.5	6.4	478	2 IS3960	PRR2 alpha - human
9	101.5	6.4	538	2 I68093	PRR2 delta - human
10	101	6.4	538	2 UC2457	vascular cell adhe
11	99	6.2	818	1 S44098	brain-derived neur
12	98	6.2	186	2 I61783	sodium channel bet
13	98	6.2	526	2 S70587	butyrophilin precu
14	97.5	6.1	315	1 HNVZVT	hemagglutinin prec
15	96	6.0	313	2 T28598	hypothetical prote
16	96	6.0	520	1 S44099	brain-derived neur
17	95	6.0	391	2 T09058	butyrophilin homol
18	95	6.0	503	2 JC5287	SHP substrate-1 pr
19	94.5	6.0	662	2 T16525	hypothetical prote
20	94.5	6.0	773	1 QRRBG	secretory componen
21	94	5.9	215	2 A57843	sodium channel bet
22	94	5.9	246	2 A47712	myelin channel bet
23	94	5.9	247	2 A55717	myelin/oligodendro
24	93	5.9	313	2 J01862	31R protein - vari
25	93	5.9	313	2 H36854	hemagglutinin - va
26	93	5.9	318	2 F72171	K9R protein - vari
27	93	5.9	372	2 C39371	Ig V-region-like B
28	92.5	5.8	309	2 I49503	B-lymphocyte activ
29	92	5.8	1091	2 S01998	contactin precurs

30	92	5.8	1134	1 JN0711	protein-tyrosine k
31	92	5.8	5175	2 T20992	hypothetical prote
32	92	5.8	5198	2 T43290	hemocentin precus
33	91.5	5.8	307	2 S55596	hypothetical prote
34	91.5	5.8	328	2 AH3342	g1px protein [limp
35	91.5	5.8	459	2 A46254	CD4 precursor - ra
36	91	5.7	1336	2 A27681	nonspecific cross-
37	91	5.7	1136	2 D6190	hypothetical prote
38	90.5	5.7	349	2 A34815	carcinoembryonic a
39	90.5	5.7	458	2 S23959	cell-adhesion mole
40	90.5	5.7	458	2 S68177	C-CAM2A protein is
41	90.5	5.7	519	2 A44783	ecto-ATPase precu
42	89.5	5.6	314	1 JQ1793	hemagglutinin prec
43	89.5	5.6	314	1 HNVZVW	hemagglutinin prec
44	89	5.6	1327	2 T09402	immunoglobulin-lik
45	88.5	5.6	136	2 I46635	rearranged T-cell

## ALIGNMENTS

## RESULT 1

S71754  
cellular hepatitis A receptor HAVcr-1 precursor - green monkey  
N/Alternate names: surface glycoprotein  
C/Species: Cercopithecus aethiops (green monkey, grivet)  
C/Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 05-Jun-1998  
C/Accession: S71754  
R/Kaplan, G.; Totsuka, A.; Thompson, P.; Akatsuka, T.; Moritsugu, Y.; Feinstein, S.M.  
EMBO J. 15, 4282-4296, 1996  
A/Title: Identification of a surface glycoprotein on African green monkey kidney cells a  
A/Reference number: S71754; MUID:97015129; PMID:8861957  
A/Accession: S71754  
A/Molecule type: mRNA  
A/Residues: 1-451 <KAP>  
A/Cross-references: EMBL:X98252; NID:G1526573; PID:6247449; PID:G1526574  
A/Experimental source: kidney  
C/Keywords: glycoprotein  
F.1-17/Domain: signal sequence #status predicted <SIG>  
F.18-451/Product: cellular hepatitis A receptor HAVcr-1 #status predicted <MAT>

Query Match	16.1%; Score 256; DB 2; Length 451;
Best Local Similarity	22.4%; Pred. No. 4e-14;
Matches	101; Conservative 38; Mismatches 123; Indels 188; Gaps 12;
QY	4 HLPFCVLLILLITRSEVEYRAVQNAVLPCTTPAAPGVNVPVCKGACPVREC 63
DB	2 HLQVILSLILHLDVSVDVAGLSTLPCRYN---GALTSCMNRKGTCSVFS 57
QY	64 GNVVLRTERDVNY-WTSRYWLNQDFPKGVSITFENVTLADSGIYCCRIQIPGIMNDEK 122
DB	58 PDGIWVINGTHTVTRKETRYKLLGNLSRRDVSLLIANTAVSDSGIYCCRVHSGFNDMK 117
QY	123 FNLKLVKIPAKV----- 134
DB	118 ITISLKIGPRVTVPIVATVSTVPTTLLPTTLLPTTLLPTTLLPTTLLPTMT 177
QY	135 ----- 134
DB	178 TLPPTTVPTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLL 237
QY	135 -----TPAPTQRFDAFPRMLT-----TRGH 157
DB	238 PTTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLL 297
QY	158 GPATQTGLSPDNLQISTLANELR---DSRIANDLRDSCATI----- 199
DB	298 EPVATSP--SSPGQAEHNPVLLGATRQPTSSSLYSITTDGSDTVSSDGMNMQ 355
QY	200 -----RIGIYIGAGICAG-LALALIFGALIFKWSHSEKTIQNLISLANLP 246
DB	356 LSPHSQPMVNTTEGIY--AGVCISVILAVLAVGVIAKKYFFKKE-IQQLS-VSFSHQ 411

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 11:56:41 ; Search time 3372.32 Seconds  
(without alignments)  
10954.308 Million cell updates/sec

Title: US-10-004-633-37

Perfect score: 903  
Sequence: 1 atgttcacacatctccctt.....gtgtgcttctgcaatgcca 903

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sfs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pac:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sfs:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rtd:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hcgo\_hum:\*  
40: em\_hcgo\_mus:\*  
41: em\_hcgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	903	100.0	903 6	AR143581
2	903	100.0	903 6	AR168950
3	903	100.0	903 6	AR232698
4	903	100.0	906 9	AY069944
5	903	100.0	1116 9	AF450242
6	903	100.0	1772 6	BD157433
7	903	100.0	1772 6	AK027334
8	903	100.0	2236 6	AR143568
9	903	100.0	2236 6	AR168937
10	903	100.0	2236 6	AR232685
11	901.4	99.8	1116 9	AF450243
12	901.4	99.8	2320 9	AF251707
13	712	78.8	769 6	BD146047
14	456	50.5	2710 6	AR143561
15	456	50.5	2710 6	AR168930
16	456	50.5	2710 6	AR232678
17	456	50.5	2710 6	189415
18	456	50.5	2725 10	AF450241
19	455	50.4	529 6	BD060055
20	442.6	49.0	843 6	AR143567
21	442.6	49.0	843 6	AR168936
22	442.6	49.0	843 6	AR232684
23	440.8	48.8	862 10	AF399831
24	396.6	43.9	1012 9	BC020843
25	374.2	41.4	438 6	BD076086
26	340.6	37.7	54375 5	AC011377
27	211.4	23.4	177087 2	AC119505
28	211.4	23.4	233370 2	AC098539
29	187.6	20.8	223030 10	AL669948
30	148	16.4	228827 2	AC135694
31	140.6	15.6	223030 10	AL669948
32	129.4	14.3	233370 2	AC098539
33	109.2	12.1	177087 2	AC119505
34	74	8.2	1507 9	BC013325
35	72.6	8.0	1440 9	AF043724
36	72.6	8.0	1795 6	BD062750
37	67.6	7.5	39980 9	AC011550
38	67.6	7.5	106169 9	AC073534
39	66.4	7.4	918 10	AF399829
40	66.4	7.4	2331 10	BC032879
41	66	7.3	140330 9	AC026777
42	66	7.3	143661 9	AC008491
43	65.4	7.2	849 10	AF399830
44	65.4	7.2	1893 10	BC03400
45	65.4	7.2	197196 2	AC112585

## ALIGNMENTS

RESULT 1  
LOCUS AR143581 903 bp DNA  
DEFINITION Sequence 37 from patent US 6204371.  
ACCESSION AR143581  
VERSION AR143581.1 GI:15104867  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 903)  
AUTHORS Levinson,D.Adam.  
TITLE Compositions and methods for the treatment and diagnosis of immune disorders  
JOURNAL Patent: US 6204371-A 37 20-MAR-2001;

FEATURES	Location/Qualifiers
Source	1..903 /organism="unknown"
BASE COUNT	242 a 225 c 221 g 215 t
ORIGIN	
Query Match	100.0%; Score 903; DB 5; Length 903;
Best Local Similarity	100.0%; Pred. No. 1,1e-277;
Matches 903; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1	ATGTTTACACATCTTCCCTTTGACGTGTGCTGCTGCTCTCTACTACTACTACAGG 60
1	ATGTTTACACATCTTCCCTTTGACGTGTGCTGCTGCTCTCTACTACTACTACAGG 60
61	TCCTCAGAAAGTGAATACAGACGGAGGTGGTGCAGAAATGCTATCTGCGCTTCTAC 120
61	TCCTCAGAAAGTGAATACAGACGGAGGTGGTGCAGAAATGCTATCTGCGCTTCTAC 120
121	ACCCAGCGCCGCCAGGGAACTCGTCCCTGCTGTGTGGGGCAAAAGAGCTGTCTGTG 180
121	ACCCAGCGCCGCCAGGGAACTCGTCCCTGCTGTGTGGGGCAAAAGAGCTGTCTGTG 180
181	TTTGAATGAGCAGGTGTGTGCTCAGACATGATGAAAGGAGTGAATATATGACATCC 240
181	TTTGAATGAGCAGGTGTGTGCTCAGACATGATGAAAGGAGTGAATATATGACATCC 240
241	AGATACCTGGCTAAATGGGGATTTCCGCAAAAGAGATGTGTCCCTGACATAGAAATGTG 300
241	AGATACCTGGCTAAATGGGGATTTCCGCAAAAGAGATGTGTCCCTGACATAGAAATGTG 300
301	ACTCTAGCAGACAGTGGGATCTTACTGTCTGCCGATCCAAATCCGACGCTAATGATATG 360
301	ACTCTAGCAGACAGTGGGATCTTACTGTCTGCCGATCCAAATCCGACGCTAATGATATG 360
361	GAATAATTTAACTCTGAAGTGTGATCAAAACGAGCAGAGTCAACCCCTGCACGACTCTG 420
361	GAATAATTTAACTCTGAAGTGTGATCAAAACGAGCAGAGTCAACCCCTGCACGACTCTG 420
421	CAGAGAGACTTCACTGACAGCTTTTCCAGAGATGCTTACCACGAGGGAGCATGCGCCAGA 480
421	CAGAGAGACTTCACTGACAGCTTTTCCAGAGATGCTTACCACGAGGGAGCATGCGCCAGA 480
481	GAGACACAGACACTGGGGAGCTCCCTGATATTAATCTTAACAACAATATCCATTGGCC 540
481	GAGACACAGACACTGGGGAGCTCCCTGATATTAATCTTAACAACAATATCCATTGGCC 540
541	AATGAGTTACGGGACTCTAGATTTGGCAATGACTTAACGGGACTCTGGACCAACATACAA 600
541	AATGAGTTACGGGACTCTAGATTTGGCAATGACTTAACGGGACTCTGGACCAACATACAA 600
601	ATAGGATCTTAATCGAGACAGGATCTGTGCTGGGCTGCTGGCTCTTATCTTCCGC 660
601	ATAGGATCTTAATCGAGACAGGATCTGTGCTGGGCTGCTGGCTCTTATCTTCCGC 660
661	GCTTTAATTTCAAAATGTTATTCATAGCAAAAGAGATACAGAAATTTAAGCTCATC 720
661	GCTTTAATTTCAAAATGTTATTCATAGCAAAAGAGATACAGAAATTTAAGCTCATC 720
721	TCCTTGGCCAACTCCCTCCCTCAGAGTTGGCAATGCAAGTACAGAGGAAATTTGGCTCA 780
721	TCCTTGGCCAACTCCCTCCCTCAGAGTTGGCAATGCAAGTACAGAGGAAATTTGGCTCA 780
781	GAAAGAAACATCTATACATTTGAAGAGATGATGAGAGTGAAGAGCCCAATGATAT 840
781	GAAAGAAACATCTATACATTTGAAGAGATGATGAGAGTGAAGAGCCCAATGATAT 840
841	TATTCATATGACAGCAGGAGCAACCTTCACAACTTTGGGTTGTGCTTTGCAATG 900
841	TATTCATATGACAGCAGGAGCAACCTTCACAACTTTGGGTTGTGCTTTGCAATG 900
901	CCA 903
901	CCA 903

RESULT 2	903 bp	DNA	Linear	PAT 17-DEC-2001
LOCUS	AR168950	903 bp	DNA	Linear
DEFINITION	Sequence 37 from patent US 6288218.			
ACCESSION	AR168950			
VERSION	AR168950.1			
KEYWORDS	GI:17905149			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 903)			
TITLE	Levinson, D. Adam.			
FEATURES	Compositions and methods for the treatment and diagnosis of immune disorders			
FEATURES	Patent: US 6288218-A 37 11-SBP-2001;			
BASE COUNT	242 a 225 c 221 g 215 t			
ORIGIN	1..903			
	/organism="unknown"			
Query Match	100.0%; Score 903; DB 6; Length 903;			
Best Local Similarity	100.0%; Pred. No. 1.1e-277; Mismatches 0; Indels 0; Gaps 0;			
Match 903; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 ATGTTTTCACATCTTCCTCTTGACTGTGTCTCTGTCTGTCTACTACTTACAGG 60			
DB	1 ATGTTTTCACATCTTCCTCTTGACTGTGTCTCTGTCTGTCTACTACTTACAGG 60			
QY	61 TCCTCAGAGTGGGAATACAGACCGGAGTGGGTGAGAAATGCTTATCTGCTCTTAC 120			
DB	61 TCCTCAGAGTGGGAATACAGACCGGAGTGGGTGAGAAATGCTTATCTGCTCTTAC 120			
QY	121 ACCCCAGCCGCCCCAGGGAACTCTGTGCGCGCTGTCTGGGGCAAGAGAGCTGTCTGTG 180			
DB	121 ACCCCAGCCGCCCCAGGGAACTCTGTGCGCGCTGTCTGGGGCAAGAGAGCTGTCTGTG 180			
QY	181 TTGTAATGTGGCAACGTGTGTCTCAGGACTGTATGAAGGATGTGAATTATTTGACATCC 240			
DB	181 TTGTAATGTGGCAACGTGTGTCTCAGGACTGTATGAAGGATGTGAATTATTTGACATCC 240			
QY	241 AGATATCTGTCTAATTTGGGGATTTTCCGAAAGGAATGTGTCTCCGACATAGAGAAATGTG 300			
DB	241 AGATATCTGTCTAATTTGGGGATTTTCCGAAAGGAATGTGTCTCCGACATAGAGAAATGTG 300			
QY	301 ACTCTAGCAGACATGGGATCTTACTGTGCGGATCCCAAAATCCAGGACATATGATGAT 360			
DB	301 ACTCTAGCAGACATGGGATCTTACTGTGCGGATCCCAAAATCCAGGACATATGATGAT 360			
QY	361 GAAAAAATTTAACTGGAAGTTGTGTCATCAACACAGGTCACCCCTGCACGACATCTGTG 420			
DB	361 GAAAAAATTTAACTGGAAGTTGTGTCATCAACACAGGTCACCCCTGCACGACATCTGTG 420			
QY	421 CAGAGAGCTTCACTGCAGACCTTTCCAAAGATGCTTACCAACAGGGGACATAGGCCACGA 480			
DB	421 CAGAGAGCTTCACTGCAGACCTTTCCAAAGATGCTTACCAACAGGGGACATAGGCCACGA 480			
QY	481 GAGACACGACACTGGGGAGCCTCCCTGATATTAATCTTAAACAAATATTCACAATTGGCC 540			
DB	481 GAGACACGACACTGGGGAGCCTCCCTGATATTAATCTTAAACAAATATTCACAATTGGCC 540			
QY	541 AATAGATTACGGGACTCTAGATTGGCCAAATGACTTAAGGGACTCTGAGGCAACATATCGA 600			
DB	541 AATAGATTACGGGACTCTAGATTGGCCAAATGACTTAAGGGACTCTGAGGCAACATATCGA 600			
QY	601 ATAGGCACTCTACATCGAGACAGGATCTGTCTGGGCTGAGCTCTTATCTTCCGGC 660			
DB	601 ATAGGCACTCTACATCGAGACAGGATCTGTCTGGGCTGAGCTCTTATCTTCCGGC 660			
QY	661 GCTTTAATTTTCAAAATGGTATTTCTCATGCAAGGAAAGATACGAATTTAAGCTTATC 720			
DB	661 GCTTTAATTTTCAAAATGGTATTTCTCATGCAAGGAAAGATACGAATTTAAGCTTATC 720			

Db	651	GCTTAAATTTTCAATVGGATTTCTCATAGCAAGAAGAAATACAGATTTTAAGCCTCATT	720
Qy	721	TCCTTTGGCCAACTCCCTTCCTCAGAGATTGGCAAAATGCAGTACAGAGGAAATTCGCTCA	780
Db	721	TCCTTTGGCCAACTCCCTTCCTCAGAGATTGGCAAAATGCAGTACAGAGGAAATTCGCTCA	780
Qy	781	GAAGAAAACATCATATACCATTTGAAGAAACGTATATGAATGTAGAGAGCCCAATGATAT	840
Db	781	GAAGAAAACATCATATACCATTTGAAGAAACGTATATGAATGTAGAGAGCCCAATGATAT	840
Qy	841	TATTGCTATGTTCAGCAGCAGCAGCAACCCCTCACAACCTTTGGTTGTGCTTTGCAATG	900
Db	841	TATTGCTATGTTCAGCAGCAGCAGCAACCCCTCACAACCTTTGGTTGTGCTTTGCAATG	900
Qy	901	CCA 903	
Db	901	CCA 903	
RESULT 3			
LOCUS	AR332698	903 bp	DNA linear PAT 20-DEC-2002
DEFINITION	Sequence 37 from patent US 6455685.		
ACCESSION	AR332698		
VERSION	AR332698.1	GI:27274975	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 903)		
AUTHORS	Levineon,D.A.		
TITLE	Compositions and methods for the treatment and diagnosis of immune disorders		
JOURNAL	Patent: US 6455685-A 37 24-SEP-2002;		
FEATURES	Location/Qualifiers		
source	1..903		
BASE COUNT	242 a 225 c 221 g 215 t		
ORIGIN			
Query Match	100.0%;	Score 903;	DB 6;
Best Local Similarity	100.0%;	Pred. No. 1.1e-277;	Length 903;
Matches	903;	Conservative 0;	Mismatches 0;
		Indels 0;	Gaps 0;
Qy	1	ATGTTTTCACATCTTCCCTTTGACTGTGTCTGCTGCTGCTGCTGCTACTACTTACAAG	60
Db	1	ATGTTTTCACATCTTCCCTTTGACTGTGTCTGCTGCTGCTGCTGCTGCTGCTACTACTTACAAG	60
Qy	61	TCCTCAGAAATGGAATACAGACGGAGGTGGTCAAGATGCTTATCTGCTGCTTTTAC	120
Db	61	TCCTCAGAAATGGAATACAGACGGAGGTGGTCAAGATGCTTATCTGCTGCTTTTAC	120
Qy	121	ACCCAGCGGCCCGAGGGGAACCTCGTCCGCTCTGCGGCAAGAGAGCCGTCTGTG	180
Db	121	ACCCAGCGGCCCGAGGGGAACCTCGTCCGCTCTGCGGCAAGAGAGCCGTCTGTG	180
Qy	181	TTTGAATGTGGCAACGTGTGCTCAGACTGATGAAAGGATGTGAATTTATTTGACATCC	240
Db	181	TTTGAATGTGGCAACGTGTGCTCAGACTGATGAAAGGATGTGAATTTATTTGACATCC	240
Qy	241	AGATATCTGGCTAAATGGGGATTTCCGCAAAAGAGATGTGTCCCTGACCAATAGAGATGTG	300
Db	241	AGATATCTGGCTAAATGGGGATTTCCGCAAAAGAGATGTGTCCCTGACCAATAGAGATGTG	300
Qy	301	ACTCAGCAGACAGTGGGATCTACTGCTGCGCGGATCCAAATGCCAGGCAATATGATATGAT	360
Db	301	ACTCAGCAGACAGTGGGATCTACTGCTGCGCGGATCCAAATGCCAGGCAATATGATATGAT	360
Qy	361	GAATAATTTAACTGTAAGTTGTGTCATCAAAACGACCAAGGTCAACCCCTGACCGACTGTG	420
Db	361	GAATAATTTAACTGTAAGTTGTGTCATCAAAACGACCAAGGTCAACCCCTGACCGACTGTG	420
Qy	421	CAGAGAGACTTACTGACGACCTTTCCAMGATGCTTACCAACGAGGACATGGCCACAGA	480

Db	421	CAGAGAGCTTCACTGCAAGCCTTTCCAGAGATGCTTACACACAGGGACATGGCCACGA	480
Qy	481	GAGACACAGACACTGGGAGGCTCCCTCGATATAATCTAACACAAATATCCATTGGCC	540
Db	481	GAGACACAGACACTGGGAGGCTCCCTCGATATAATCTAACACAAATATCCATTGGCC	540
Qy	541	AATGATGACGGGACTCTAGATTGGCCCAATGACTTTACGGGACTCTGGAGCAACATCAGA	600
Db	541	AATGATGACGGGACTCTAGATTGGCCCAATGACTTTACGGGACTCTGGAGCAACATCAGA	600
Qy	601	ATAGGCAATCTACATCGGAGCAGGGGATCTGTGCTGTGGCTGTGCTTATCTTGGCC	660
Db	601	ATAGGCAATCTACATCGGAGCAGGGGATCTGTGCTGTGGCTGTGCTTATCTTGGCC	660
Qy	661	GCTTTAATTTTCAATGGTATTTCTCATAGCAAGAAGATACAGAAATTTAAGCTCATC	720
Db	661	GCTTTAATTTTCAATGGTATTTCTCATAGCAAGAAGATACAGAAATTTAAGCTCATC	720
Qy	721	TCCTTTGGCAACCTCCCTCCCTCAGAGATTGGCAATGACATGACGAGAGGGAATTCGCTCA	780
Db	721	TCCTTTGGCAACCTCCCTCCCTCAGAGATTGGCAATGACATGACGAGAGGGAATTCGCTCA	780
Qy	781	GAAAGAAAACATCTATTACCTTTGAAGAGAAAGTAAATGAAGTCAGAGAGCCCAATGAGTAT	840
Db	781	GAAAGAAAACATCTATTACCTTTGAAGAGAAAGTAAATGAAGTCAGAGAGCCCAATGAGTAT	840
Qy	841	TATTGCTATGTACAGACAGGACAGCAACCTTCACAACTTTGGGTTGTGCTTTGCATG	900
Db	841	TATTGCTATGTACAGACAGGACAGCAACCTTCACAACTTTGGGTTGTGCTTTGCATG	900
Qy	901	CCA 903	
Db	901	CCA 903	
RESULT 4			
LOCUS	AY069944	906 bp	mRNA linear PRI 01-JUL-2002
DEFINITION	Homo sapiens putative kidney injury molecule-3 mRNA, complete cds.		
ACCESSION	AY069944		
VERSION	AY069944.1	GI:21655105	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Kuehn, E.W., Ichimura, T. and Bonventre, J.V.		
TITLE	A homolog to human kidney injury molecule-1 is expressed in hepatoma cells		
JOURNAL	unpublished		
REFERENCE	2 (bases 1 to 906)		
AUTHORS	Kuehn, E.W., Ichimura, T. and Bonventre, J.V.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-DEC-2001) Renal Unit, Massachusetts General Hospital, 149, 13th Street, Charlestown, MA 02129, USA		
FEATURES	Location/Qualifiers		
source	1..906		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="5"		
	/cell_line="HepG2"		
	/tissue_type="hepatocellular carcinoma"		
	1..906		
	/note="hKIM-3; similar to Homo sapiens kidney injury molecule-1 (hKIM-1)"		
	/codon_start=1		
	/product="putative kidney injury molecule-3"		
	/protein_id="AA55401.1"		
	/db_xref="GI:21655106"		
	/translation="MFSLHPFDVCLLLLLLRSSSEVRAEIVGQNAVYLPCTYPA		

[illegible]

DB	901	CCA	903
RESULT 5	AF450242	1116 bp	mRNA
LOCUS	AF450242		linear
DEFINITION	Homo sapiens clone 1 T cell immunoglobulin mucin-3 (TIM3) mRNA,		PRI 11-FEB-2002
ACCESSION	AF450242		
VERSION	AF450242.1		
KEYWORDS	GI:18182532		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1116) Monney, L., Sabatos, C.A., Gaglia, J.L., Ryu, A., Waldner, H., Freeman, G.J., Manning, S., Greenfield, E.A., Coyle, A.J., Sobel, R.A., Chernova, T., and Kuchroo, V.K. Tim-3-specific cell surface protein Tim-3 regulates macrophage activation and severity of an autoimmune disease Nature 415 (6871), 536-541 (2002)		
TITLE	2 (bases 1 to 1116) Monney, L., Sabatos, C., Gaglia, J.L., Ryu, A., Waldner, H., Chernova, T., Greenfield, E.A., Sobel, R.A., Freeman, G.J. and Kuchroo, V.K. Direct Submission		
JOURNAL	Submitted (23-NOV-2001) Adult Oncology, Dana-Farber Cancer Institute, 44 Binney St., Boston, MA 02115, USA		
PUBMED	Location/Qualifiers		
REFERENCE	1. 1116 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="1" /cell_type="T cell" 1. 1116 /gene="TIM3" 58. 963 /gene="TIM3" /codon_start=1 /product="T cell immunoglobulin mucin-3" /protein_id="AA65157.1" /db_xref="GI:18182533" /translation="MFSHLPFDDCYLLLLLTTRSSVEYRAEYQNAVYLPFFYTPAA PGNIVPOMGKACPVCEGVVAVRTDERDVNVTSYWLNGDPKQSVSLTINVTI ADSGYICCRIDIPGIMNDEKNIKLVIKPKVPAPLQRFPTAPRMLTRHGSP ETQYLFSPDNLQIOTSLANELRDSPLANDRSGATRIIGYIGICGALALIL FGALIFMYSHSKETQNLISLANLPISGLANAVABGRSENIYTIENVTVEE PNEYCVVSSHQPSQPLGCRFAMP"		
AUTHORS	BASE COUNT 291 a 268 c 270 g 287 t		
TITLE	ORIGIN		
JOURNAL	Query Match 100.0%; Score 903; DB 9; Length 1116; Best Local Similarity 100.0%; Pred. No. 1.1e-277; Indels 0; Gaps 0; Matches 903; Conservative 0; Mismatches 0;		
AUTHORS	1 ATGTTTACATCTCTCTTGACTGTGCTCTGCTGCTGCTGCTACTACTTACAAAGG 60		
JOURNAL	58 ATGTTTACATCTCTCTTGACTGTGCTCTGCTGCTGCTGCTGCTACTACTTACAAAGG 117		
PUBMED	61 TCCTCAGAAGTGGAATACAGACGGAGGTCGGTCAGAAATGCTTATCTGCGCTTCTAC 120		
REFERENCE	118 TCCTCAGAAGTGGAATACAGACGGAGGTCGGTCAGAAATGCTTATCTGCGCTTCTAC 177		
AUTHORS	121 ACCCAGCGCGCCCAAGGGAACCTGCTGCGCTGCTGCGGCAAGAGAGCCTGCTCTGTG 180		
TITLE	178 ACCCAGCGCGCCCAAGGGAACCTGCTGCGCTGCTGCGGCAAGAGAGCCTGCTCTGTG 237		
JOURNAL	181 TTGAGATGTGGCAAGTGTGCTCAGACATGATGAAGAGATGGAATTAATTCGACATCC 240		

Db	238	TTGATATGTGGCAACGTGTGTCTCAGACTGATGAAGGATGTGAATTTATTGGACATCC	297
Qy	241	AGATATCTGGCTAAATGGGGATTTTCCGCAAGAGATGTGTCTCTGACCATAGAGAAATGTG	300
Db	298	AGATATCTGGCTAAATGGGGATTTTCCGCAAGAGATGTGTCTCTGACCATAGAGAAATGTG	357
Qy	301	ACTCTAGAGACAGTGGGATCTACCTGCTCGCCGGAATCCAAATCCAGGGCATTAATGAAATGAT	360
Db	358	ACTCTAGAGACAGTGGGATCTACCTGCTCGCCGGAATCCAAATCCAGGGCATTAATGAAATGAT	417
Qy	361	GAATAATTTAACTTGAAGTTGTGATCAATCAACCAAGCAAGTCAACCCCTGCACCGACTGTG	420
Db	418	GAATAATTTAACTTGAAGTTGTGATCAATCAACCAAGCAAGTCAACCCCTGCACCGACTGTG	477
Qy	421	CAGAGAGCTTCACTGCGACCTTTTCCAGAGATGCTTACCACAGGGGACATATGGCCACGA	480
Db	478	CAGAGAGCTTCACTGCGACCTTTTCCAGAGATGCTTACCACAGGGGACATATGGCCACGA	537
Qy	481	GAGACACAGACACTGGGGAGCCTCCCTGATATTAATCTAACCAAAATATCCACATTGGCC	540
Db	538	GAGACACAGACACTGGGGAGCCTCCCTGATATTAATCTAACCAAAATATCCACATTGGCC	597
Qy	541	AATGATTTACGGGACTCTAGATTGTGCAATGACTTACGGGACTCTGAGCAACATCAGA	600
Db	598	AATGATTTACGGGACTCTAGATTGTGCAATGACTTACGGGACTCTGAGCAACATCAGA	657
Qy	601	ATAGGCATCTACATCGGAGCAAGGAGATCTGTGCTGGGCTGTGGCTCTTATCTTTGGG	660
Db	658	ATAGGCATCTACATCGGAGCAAGGAGATCTGTGCTGGGCTGTGGCTCTTATCTTTGGG	717
Qy	661	GCTTTAATTTTCAATGTGTATTTCTATAGCAAAAGAAATACAGAAATTTAAAGCTCATC	720
Db	718	GCTTTAATTTTCAATGTGTATTTCTATAGCAAAAGAAATACAGAAATTTAAAGCTCATC	777
Qy	721	TCTTTGGCAACCTCCCTCCCTCAGAGATTGGCAATGCAATGACAGAGGAATTTGGCTCA	780
Db	778	TCTTTGGCAACCTCCCTCCCTCAGAGATTGGCAATGCAATGACAGAGGAATTTGGCTCA	837
Qy	781	GAAAGAAAACATCTATATACCATTTGAAGAGACGTAATGAAGTGGAGAGAGCCCAATGAGTAT	840
Db	838	GAAAGAAAACATCTATATACCATTTGAAGAGACGTAATGAAGTGGAGAGAGCCCAATGAGTAT	897
Qy	841	TATTGCTATGTACAGCAGCAGCAGCAACCTTCACACCTTTGGGTGTGTGCTTTGCAATG	900
Db	898	TATTGCTATGTACAGCAGCAGCAGCAGCAACCTTCACACCTTTGGGTGTGTGCTTTGCAATG	957
Qy	901	CCA 903	
Db	958	CCA 960	
RESULT 6			
BD157433			
LOCUS	BD157433	1772 bp	DNA linear
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.		RAF 17-JAN-2002
ACCESSION	BD157433		
VERSION	BD157433.1	GI:27863191	
KEYWORDS	JP 2002191363-A/12276		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 1772)		
JOURNAL	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.		
COMMENT	Primer for synthesizing full-length cDNA and use thereof		
	HELEX RESEARCH INSTITUTE		
	OS Homo sapiens (human)		
	PN JP 2002191363-A/12276		
	PD 09-JUL-2000		
	PF 28-JUL-2000 JP 2000280990		

FEATURES	source
PI	TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI	SAITO,
PI	JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU
PI	KEIICHI NAKAI, TETSUJI OTSUKI
PC	
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10,	
PC	C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N5/00, C12N5/00 CC
Primer	for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers	
FT	CDS (44)..(946).
	Location/Qualifiers
	1..11772
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
BASE COUNT	473 a 396 c 419 g 484 t
ORIGIN	

Query Match	100.0%;	Score 903;	DB 6;	Length 1772;
Best Local Similarity	100.0%;	Pred. No. 1.2e-277;		
Matches 903;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Oy	1	ATGTTTACATCTTCCCTTGA	CTGTGTCTGTGCTGTCTCTACTTACCAAG	60
Db	44	ATGTTTCAATCTTCCCTTGA	CTGTGTCTGTGCTGTCTCTACTTACCAAG	103
Oy	61	TCCTCAGAACTGSAATA	CAGAGCGAGGTGGTCAGAAATGCTTATCTGCTCTTCTAC	120
Db	104	TCCTCAGAAAGTGAAAT	CAGAGCGAGGTGGTCAGAAATGCTTATCTGCTCTTCTAC	163
Oy	121	ACCCCAACCGCCCAAGGAA	ACTGTGTGCTGTCTGTGCGGCAAAAGAGCTGTCTGTG	180
Db	164	ACCCCAACCGCCCAAGGAA	ACTGTGTGCTGTCTGTGCGGCAAAAGAGCTGTCTGTG	223
Oy	181	TTTGAATGTGGCAAGTGT	GTCTCAGGACTATGAAGAATGTGAATTAATTTGACATCC	240
Db	224	TTTGAATGTGGCAAGTGT	GTCTCAGGACTATGAAGAATGTGAATTAATTTGACATCC	283
Oy	241	AGATACTGGCTMAATGG	GGATTTTCCGCAAGAGAGTGTGCTCCCTGACCATAGAGATGTG	300
Db	284	AGATACTGGCTMAATGG	GGATTTTCCGCAAGAGAGTGTGCTCCCTGACCATAGAGATGTG	343
Oy	301	ACTTACGACAGATGGGAT	TACTCTGTCCGGATCCAAATCCACAGCATTAATGATAT	360
Db	344	ACTTACGACAGATGGGAT	TACTCTGTCCGGATCCAAATCCACAGCATTAATGATAT	403
Oy	361	GAAAAATTTAACCTGAA	ATGGTGTGATCAAAACCAACGAAAGTCAACCCCTGCACCGGACTGTG	420
Db	404	GAAAAATTTAACCTGAA	ATGGTGTGATCAAAACCAACGAAAGTCAACCCCTGCACCGGACTGTG	463
Oy	421	CAGAGACCTTCATGCA	GCTTTTCCAAAGATGCTTATCACACAGGGGACATAGCCACACA	480
Db	464	CAGAGACCTTCATGCA	GCTTTTCCAAAGATGCTTATCACACAGGGGACATAGCCACACA	523
Oy	481	GAGACACAGACACTGG	GGAGGCTCCCTGTATTAATCTTAACACAATAATTCACATTTGGCC	540
Db	524	GAGACACAGACACTGG	GGAGGCTCCCTGTATTAATCTTAACACAATAATTCACATTTGGCC	583
Oy	541	AATAGTTTACGGGACT	CTAGATTGGCCATTACTTAACGGGACTCTGGAGCAACATTCAGA	600
Db	584	AATAGTTTACGGGACT	CTAGATTGGCCATTACTTAACGGGACTCTGGAGCAACATTCAGA	643
Oy	601	ATAGGCAATCAATCGG	AGCAGGAGCTGTGCTGGGCTGGCTCTGAGCTTATCTTCGGC	660
Db	644	ATAGGCAATCAATCGG	AGCAGGAGCTGTGCTGGGCTGGCTCTGAGCTTATCTTCGGC	703
Oy	661	GCTTAAATTTCAAAAT	GGTATTTCTCATAGCAAAAGAGATACAGAAATTTAAGCTCATC	720
Db	704	GCTTAAATTTCAAAAT	GGTATTTCTCATAGCAAAAGAGATACAGAAATTTAAGCTCATC	763
Oy	721	TCTTTGGCAACTTCCT	CTCCTCAGATTTGGCAAAATGACATAGCAGAGGAAATTCCTCA	780

DB	764	1CTTTGGCAACCTCCCTCCCTCAGAGATTGGGCAAAATGACATGACAGAGAGAAATTCGCTCA	823
QY	781	GAGAAAAACATCTATACCATTTGAAGAGAACGTAATATGAATGAGAGAGAGAGAGAGAT	840
DB	824	GAGAAAAACATCTATACCATTTGAAGAGAACGTAATATGAATGAGAGAGAGAGAGAT	883
QY	841	TATGCTATGTCAGAT	900
DB	884	TATGCTATGTCAGAT	943
QY	901	CCA 903	
DB	944	CCA 946	
RESULT 7	AK027334	1772 bp	mRNA linear PRI 01-NOV-2002
LOCUS	AK027334		
DEFINITION	Human sapiens CDNA FLJ14428 fis, clone HEMBA1006293.		
ACCESSION	AK027334		
VERSION	AK027334.1	GI:14041941	
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Human sapiens (human)		
ORGANISM	Human sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1		
TITLE	Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Nabekura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, T., Wakamatsu, A., Nakamura, Y., Nagahashi, K., Masuho, Y. and Oshima, A.		
JOURNAL	NEBD human cDNA sequencing project		
REFERENCE	2 (bases 1 to 1772)		
AUTHORS	Isogai, T. and Otsuki, T.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)		
COMMENT	NEBD human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.		
FEATURES	Location/Qualifiers		
source	1..1772		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="HEMBA1006293"		
	/tissue_type="whole embryo, mainly head"		
	/clone_lib="HEMBA1"		
	/dev_stage="embryo, 10 weeks"		
	/note="Cloning vector: pME18SFL3"		
	44..949		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="BAB55044.1"		
	/db_xref="GI:14041942"		
	/translation="MFSLPFCVLLLLLLLTSSSEVEYRAEYGVNAYLPFCPTTPAA		
	PENLVPCMGKACVPEFGKVLATDERDVNWTSMWNGDPRKGVSLTINVTLL		
	ADSGVYVCRCIOIPGIMNDEKNKLVKIVAPKATVAPITCRDFTAAFPMTLTTRHGA		
	EYQTGSLPDINLTQISTLANELRKRSLANDLRSGATIRIGYITGAGICAGLALALI		
	FEALFKMYSHSKETQNLISLNLDPGGLANNAVAGIRSEENITTEENYVEVEE		
	PENLYCYVSSRQOPQPGCRFAMP"		
BASE COUNT	473 a 396 c 419 g 484 t		
ORIGIN			
Query Match	100.0%; Score 903; DB 9; Length 1772;		
Best Local Similarity	100.0%; Pred. No. 1.2e-277;		

	Matches	903: Conservative	0: Mismatches	0: Indels	0: Gaps	0:
QY	1	ATGTTTTCACATCTTCCCTTTGA	CTGTGCTCTGCTGCTACTTAC	CAAG	60	
Db	44	ATGTTTTCACATCTTCCCTTTGA	CTGTGCTCTGCTGCTACTTAC	CAAG	103	
QY	61	TCCTCACAATGGAATCAGAGCG	AGGTGCTCAGAAATCCATATC	CGCTTCTAC	120	
Db	104	TCCTCACAATGGAATCAGAGCG	AGGTGCTCAGAAATCCATATC	CGCTTCTAC	163	
QY	121	ACCCCAAGCCGCCCCAGGGA	ACTCGTCCCGTCTGCTGGGG	CAAAAGAGCCTGCTGTG	180	
Db	164	ACCCCAAGCCGCCCCAGGGA	ACTCGTCCCGTCTGCTGGGG	CAAAAGAGCCTGCTGTG	223	
QY	181	TTTGAATGTGGCAACGTGTG	CTCTCAGCATGTAAAGGATG	TAATTAATGGACATCC	240	
Db	224	TTTGAATGTGGCAACGTGTG	CTCTCAGCATGTAAAGGATG	TAATTAATGGACATCC	283	
QY	241	AGATATCTGGCTAAATGGGG	ATTTCCGAAAGAGATGTGTCC	TGACCATAGAAATGTG	300	
Db	284	AGATATCTGGCTAAATGGGG	ATTTCCGAAAGAGATGTGTCC	TGACCATAGAAATGTG	343	
QY	301	ACTTTAGCAGACAGTGGGAT	CTACTGTGCGGATCCAAATCC	CAGGCATATGAATGAT	360	
Db	344	ACTTTAGCAGACAGTGGGAT	CTACTGTGCGGATCCAAATCC	CAGGCATATGAATGAT	403	
QY	361	GAATAATTTAACCCTGAAG	TGTCATCAACACAGCCAAAG	TCAACCCCTGACCGACTGTG	420	
Db	404	GAATAATTTAACCCTGAAG	TGTCATCAACACAGCCAAAG	TCAACCCCTGACCGACTGTG	463	
QY	421	CAGAGACATTTCACTGACG	CTTTCCCAAGATGTTTCAAC	ACAGGGACATATGGCCACGA	480	
Db	464	CAGAGACATTTCACTGACG	CTTTCCCAAGATGTTTCAAC	ACAGGGACATATGGCCACGA	523	
QY	481	GAGACACAGACACTGGGGAG	CCCTCGATATATAATCTA	CACAAAATATCCACTTGGCC	540	
Db	524	GAGACACAGACACTGGGGAG	CCCTCGATATATAATCTA	CACAAAATATCCACTTGGCC	583	
QY	541	AATGAGTTTACGGGACTCT	AGATTGGCCAATGACTTAC	GGGACTCTGAGCAACCATCAGA	600	
Db	584	AATGAGTTTACGGGACTCT	AGATTGGCCAATGACTTAC	GGGACTCTGAGCAACCATCAGA	643	
QY	601	ATTAGGATCTTACATCGGA	GCAGGGATCTGTGCTGAGG	CTGTGCTCTTATCTTGGGC	660	
Db	644	ATTAGGATCTTACATCGGA	GCAGGGATCTGTGCTGAGG	CTGTGCTCTTATCTTGGGC	703	
QY	661	GCTTTAATTTTCAAAATG	TATTCATAGCAAAAGAGA	ATPACGAATTTAAGCTCATC	720	
Db	704	GCTTTAATTTTCAAAATG	TATTCATAGCAAAAGAGA	ATPACGAATTTAAGCTCATC	763	
QY	721	TCCTTTGGCCAACTCCCT	CCCTCAGATTGSCAAATC	CAGTAGCAGAGGAAATTCGCTCA	780	
Db	764	TCCTTTGGCCAACTCCCT	CCCTCAGATTGSCAAATC	CAGTAGCAGAGGAAATTCGCTCA	823	
QY	781	GAAAGAAAACATCTATAC	ATTGAAGAGACGTATATGA	AGTGAAGAGGCCCAATGAATAT	840	
Db	824	GAAAGAAAACATCTATAC	ATTGAAGAGACGTATATGA	AGTGAAGAGGCCCAATGAATAT	883	
QY	841	TATTCGATATGACACAC	AGGCAACCTCTCAAACTTT	GGGTGTGCTTGGCTTCATG	900	
Db	884	TATTCGATATGACACAC	AGGCAACCTCTCAAACTTT	GGGTGTGCTTGGCTTCATG	943	
QY	901	CCA 903				
Db	944	CCA 946				
RESULT 8						
LOCUS	ARI43568	2236 bp	DNA	linear	PAT 08-AUG-2001	
DEFINITION	Sequence 23 from patent US 6204371.					
ACCESSION	ARI43568					
VERSION	ARI43568.1	GI:15104854				





```

Db 582 AATGATGAGGAGCTGATGATGCGAATGACTTACGAGCTTGAGACATCAGA 641
Qy 601 ATAGCATCTACATCGAGCAGGAGATCTGTGCTGAGCTTGCTTATCTTGCG 660
Db 642 ATAGCATCTACATCGAGCAGGAGATCTGTGCTGAGCTTGCTTATCTTGCG 701
Qy 661 GCTTTAATTTTCAATGATGATTTCTCATAGCAAGAGATACAGAAATTTAGCTTCATC 720
Db 702 GCTTTAATTTTCAATGATGATTTCTCATAGCAAGAGATACAGAAATTTAGCTTCATC 761
Qy 721 TCTTTGGCAACCTCCCTCCCTCAGAGTTGGCAAAATGCACTAGCAGGGGAATTTGCTCA 780
Db 762 TCTTTGGCAACCTCCCTCCCTCAGAGTTGGCAAAATGCACTAGCAGGGGAATTTGCTCA 821
Qy 781 GAAGAAACATCTATACCATTTGAAGAGAGATATGATGAGAGAGCCCATGAGTAT 840
Db 822 GAAGAAACATCTATACCATTTGAAGAGAGATATGATGAGAGAGCCCATGAGTAT 881
Qy 841 TATTGCTATGTGAGCAGCAGGAGCAACCTTCACAACTTTGGGTTGTGCTTTGCAATG 900
Db 882 TATTGCTATGTGAGCAGCAGGAGCAACCTTCACAACTTTGGGTTGTGCTTTGCAATG 941
Qy 901 CCA 903
Db 942 CCA 944

RESULT 10
AR232685 2236 bp DNA linear PAT 20-DEC-2002
LOCUS AR232685
DEFINITION Sequence 23 from patent US 6455685.
ACCESSION AR232685
VERSION AR232685.1 GI:27274962
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2236)
AUTHORS Levinson,D.A.
TITLE Compositions and methods for the treatment and diagnosis of immune disorders
JOURNAL Patent: US 6455685-A 23 24-SEP-2002;
FEATURES
Source 1. 2236
Location/Qualifiers
BASE COUNT 586 a 518 c 525 g 606 t 1 others
ORIGIN
Query Match 100.0%; Score 903; DB 6; Length 2236;
Best Local Similarity 100.0%; Pred.No.1.2e-277;
Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 301 ACTCTAGCAGACAGTGGGATCTACCTGCTCCGGATCCAAATCCAGCATATGATAT 360
Db 342 ACTCTAGCAGACAGTGGGATCTACCTGCTCCGGATCCAAATCCAGCATATGATAT 401
Qy 361 GAAATATTTAACTCAAGTTGGTTCATCAAAACGAGCAGATGATCCCTTGACCGACTCG 420
Db 402 GAAATATTTAACTCAAGTTGGTTCATCAAAACGAGCAGATGATCCCTTGACCGACTCG 461
Qy 421 CAGAGACCTTCACTGACGCTTTCCAAAGATGCTTACCAACGAGGAGCATGGCCAGCA 480
Db 462 CAGAGACCTTCACTGACGCTTTCCAAAGATGCTTACCAACGAGGAGCATGGCCAGCA 521
Qy 481 GAGACAGACACCTGGGAGCTCCCTGATATTAATCAACAATAATCCATTTGGCC 540
Db 522 GAGACAGACACCTGGGAGCTCCCTGATATTAATCAACAATAATCCATTTGGCC 581
Qy 541 AATGATTTACGGAGCTTCAATTTGGCAATGACTTACGAGACTTGGAGCAACATCAGA 600
Db 582 AATGATTTACGGAGCTTCAATTTGGCAATGACTTACGAGACTTGGAGCAACATCAGA 641
Qy 601 ATAGCATCTACATCGAGCAGGAGATCTGTGCTGAGCTTGCTTATCTTGCGC 660
Db 642 ATAGCATCTACATCGAGCAGGAGATCTGTGCTGAGCTTGCTTATCTTGCGC 701
Qy 661 GCTTTAATTTTCAATGATGATTTCTCATAGCAAGAGATACAGAAATTTAGCTTCATC 720
Db 702 GCTTTAATTTTCAATGATGATTTCTCATAGCAAGAGATACAGAAATTTAGCTTCATC 761
Qy 721 TCTTTGGCAACCTCCCTCCCTCAGAGTTGGCAAAATGCACTAGCAGGGGAATTTGCTCA 780
Db 762 TCTTTGGCAACCTCCCTCCCTCAGAGTTGGCAAAATGCACTAGCAGGGGAATTTGCTCA 821
Qy 781 GAAGAAACATCTATACCATTTGAAGAGAGATATGATGAGAGAGCCCATGAGTAT 840
Db 822 GAAGAAACATCTATACCATTTGAAGAGAGATATGATGAGAGAGCCCATGAGTAT 881
Qy 841 TATTGCTATGTGAGCAGCAGGAGCAACCTTCACAACTTTGGGTTGTGCTTTGCAATG 900
Db 882 TATTGCTATGTGAGCAGCAGGAGCAACCTTCACAACTTTGGGTTGTGCTTTGCAATG 941
Qy 901 CCA 903
Db 942 CCA 944

RESULT 11
AF450243 1116 bp mRNA linear PRI 11-FEB-2002
LOCUS AF450243
DEFINITION Homo sapiens clone 2 T cell immunoglobulin mucin-3 (TIM3) mRNA, complete cds.
ACCESSION AF450243
VERSION AF450243.1 GI:18182534
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1116)
AUTHORS Mammlaia; Eutheia; Primates; Catherini; Homindae; Homo.
TITLE Th1-specific cell surface protein Tim-3 regulates macrophage activation and severity of an autoimmune disease
JOURNAL Nature 415 (6871), 536-541 (2002)
MEDLINE 21681888
PUBMED 11823861
REFERENCES
1 (bases 1 to 1116)
Mammlaia, L., Sabatos, C.A., Gaglia, J.L., Ryu, A., Waldner, H., Chernova, T., Manning, S., Greenfield, E.A., Coyle, A.J., Sobel, R.A., Freeman, G.J. and Kuchroo, V.K.
Th1-specific cell surface protein Tim-3 regulates macrophage activation and severity of an autoimmune disease
Submitted (23-NOV-2001) Adult Oncology, Dana-Farber Cancer Institute, 44 Binney St., Boston, MA 02115, USA

```



Qy	121	ACCACGACGCGCCACGAGGAACCTGTGTCGCTGTGAGGACAAAGAGCCTGTG	180
Db	247	ACCACGACGCGCCACGAGGAACCTGTGTCGCTGTGAGGACAAAGAGCCTGTG	306
Qy	181	TTTGAATGTGGCAACGTGTGCTCAGACCTGATGAAGAGTGTGAATTATTTGACATCC	240
Db	307	TTTGAATGTGGCAACGTGTGCTCAGACCTGATGAAGAGTGTGAATTATTTGACATCC	366
Qy	241	AGATATCTGTGCTAAATGTGGGATTTTCCGCAAGAGATGTGTCCCTGACATAGAGATGTG	300
Db	367	AGATATCTGTGCTAAATGTGGGATTTTCCGCAAGAGATGTGTCCCTGACATAGAGATGTG	426
Qy	301	ACTCTAGACGACAGTGGGATCTACCTGCTGCGGATCCAAATCCGACGATTAATGAT	360
Db	427	ATTCTAGACGACAGTGGGATCTACCTGCTGCGGATCCAAATCCGACGATTAATGAT	486
Qy	361	GAATAATTTAACTGAAAGTGTGTCATCAACACAGCCAAAGTCACCCCTGCACCGACTCTG	420
Db	487	GAATAATTTAACTGAAAGTGTGTCATCAACACAGCCAAAGTCACCCCTGCACCGACTCTG	546
Qy	421	CAGAGAGCTTCACTGCAGCCTTTTCCAAAGATGCTTACACACAGGGGACATATGGCCACGA	480
Db	547	CAGAGAGCTTCACTGCAGCCTTTTCCAAAGATGCTTACACACAGGGGACATATGGCCACGA	606
Qy	481	GAGACACAGACACTGGGGAGCCTCCCTGATATTAAATCTAACAAATATCCACTTGGACC	540
Db	607	GAGACACAGACACTGGGGAGCCTCCCTGATATTAAATCTAACAAATATCCACTTGGACC	666
Qy	541	AATGATTTACGGGACTCTGATTGGCCAAATGACTTACCGGACTCTGAGCAACCATGAGA	600
Db	667	AATGATTTACGGGACTCTGATTGGCCAAATGACTTACCGGACTCTGAGCAACCATGAGA	726
Qy	601	ATAGGCATCTACATCGAGCAAGGATCTGTGCTGGGCTGTGGCTTATCTTGCGC	660
Db	727	ATAGGCATCTACATCGAGCAAGGATCTGTGCTGGGCTGTGGCTTATCTTGCGC	786
Qy	661	GCTTAAATTTTCAATGTGTTATCTCATAGCAAGAGAAATTAATTAAGCCTCATC	720
Db	787	GCTTAAATTTTCAATGTGTTATCTCATAGCAAGAGAAATTAATTAAGCCTCATC	846
Qy	721	TCTTTGGCCAACTCTCCTCCTCAGAGATTGGCAAAATGACATGACAGAGGGAATTCGCTCA	780
Db	847	TCTTTGGCCAACTCTCCTCCTCAGAGATTGGCAAAATGACATGACAGAGGGAATTCGCTCA	906
Qy	781	GAGAGAAACATCTTATACCATTTGAAGAGAACGTATATGAAGTGGAGAGGCCCAATGAGAT	840
Db	907	GAGAGAAACATCTTATACCATTTGAAGAGAACGTATATGAAGTGGAGAGGCCCAATGAGAT	966
Qy	841	TATTGCTATGTACAGCAGCAGGACAGCAACCTTCACAACTTTGGGTGTGCTTTTGCATG	900
Db	967	TATTGCTATGTACAGCAGCAGGACAGCAACCTTCACAACTTTGGGTGTGCTTTTGCATG	1026
Qy	901	CCA 903	
Db	1027	CCA 1029	
RESULT 13			
BD146047			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			

JOURNAL	Patent: JP 2002191363-A 890 09-JUL-2002;									
COMMENT	HELIIX RESEARCH INSTITUTE									
OS	Homo sapiens (human)									
PN	JP 2002191363-A/890									
PD	09-JUL-2002									
PF	28-JUL-2000 JP 2000280990									
P1	TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO,									
P1	JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU									
P1	KEIICHI NAGAI, TETSUJI OTSUKI.									
PC	C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10,									
PC	C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC									
Primer	for synthesizing full-length cDNA and use thereof FH Key									
Location/Qualifiers										
FT	source	1..769								
FT	Location/Qualifiers	/organism='Homo sapiens (human)'								
source	/organism='Homo sapiens'									
	/mol_type='Genomic DNA'									
	/db_xref='taxon:9606'									
BASE COUNT	195 a 194 c 189 g 188 t 3 others									
ORIGIN										
Query Match	78.8%; Score 712; DB 6; Length 769;									
Best Local Similarity	99.4%; Pred. No. 1.9e-216;									
Matches	723; Conservative 0; Mismatches 3; Indels 1; Gaps 1									
QY	1	ANGTTTTCACATCTTCCCTTTGACTGTGTCTCTGCTGCTGCTGCTACTACTTACAAG	60							
DB	44	ANGTTTTCACATCTTCCCTTTGACTGTGTCTCTGCTGCTGCTGCTACTACTTACAAG	103							
QY	61	TCCTCAGAGTGGAAATACAGAGCGAGGTGCGTCAGATGCTATCTGCCTCTTCTAC	120							
DB	104	TCTTCGAAGTGGATACAGAGCGAGGTGCGTCAGATGCTATCTGCCTCTTCTAC	163							
QY	121	ACCCGACGCGCCCAAGGAACTGTGTGCTGTGCGGCAAAAGAGCTGTCTGTG	180							
DB	164	ACCCGACGCGCCCAAGGAACTGTGTGCTGTGCGGCAAAAGAGCTGTCTGTG	223							
QY	181	TTTGAATGTGGCAAGTGTGTCTACAGACTGATGAAAGGATGTGAATTTTGGACATCC	240							
DB	224	TTTGAATGTGGCAAGTGTGTCTACAGACTGATGAAAGGATGTGAATTTTGGACATCC	283							
QY	241	AGATCTGCTAAATGGGGATTTCCGCCAAAGAGATGTGTCCCTGACCATAGAAATGTG	300							
DB	284	AGATCTGCTAAATGGGGATTTCCGCCAAAGAGATGTGTCCCTGACCATAGAAATGTG	343							
QY	301	ACCTTAGCAGAGTGGGATCTACGTGCTGCCGGAATCCAAATCCGAGCATTAAGATGAT	360							
DB	344	ACCTTAGCAGAGTGGGATCTACGTGCTGCCGGAATCCAAATCCGAGCATTAAGATGAT	403							
QY	361	GAATAATTTAACTGAAGTTGTGTCATCAACACAGCCAAAGTCAACCCCTGACCGACTTG	420							
DB	404	GAATAATTTAACTGAAGTTGTGTCATCAACACAGCCAAAGTCAACCCCTGACCGACTTG	463							
QY	421	CAGAGGACTTCACTGCAAGCTTTTCAAGATGCTTACACACAGGGGACATGGCCACGCA	480							
DB	464	CAGAGGACTTCACTGCAAGCTTTTCAAGATGCTTACACACAGGGGACATGGCCACGCA	523							
QY	481	GAGACACAGACCTGGGGGAGCTCCCTGATTTAATATCTAACACAAATATCCATTGGCC	540							
DB	524	GAGACACAGACCTGGGGGAGCTCCCTGATTTAATATCTAACACAAATATCCATTGGCC	583							
QY	541	AATGAGTTACGGGACTTAGATTGGCCAAATGCTTACCGGACTCTGAGCAACCATCAGA	600							
DB	584	AATGAGTTACGGGACTTAGATTGGCCAAATGCTTACCGGACTCTGAGCAACCATCAGA	643							
QY	601	ATAGGATCTTAATGTGGACAGGAGTCTGTGCTGGGCTGGCTGGCTCTTATCTTCGGC	660							
DB	644	ATAGGATCTTAATGTGGACAGGAGTCTGTGCTGGGCTGGCTGGCTCTTATCTTCGGC	703							



```
Db 340 G|T|A|C|T|G|G|A|T|C|C|A|T|G|G|A|C|T|A|C|T|G|C|A|G|A|T|C|A|G|T|T|C|C|T|G|T|T|A|T|A|T| 399
QY 358 G|A|T|G|A|A|A|A|T|T|T|A|C|C|T|G|A|A|G|T|G|T|C|A|T|C|A|A|C|C|A|G|G|T|C|A|C|C|C|T|G|C|A|C|C|A|C|T| 417
Db 400 G|A|T|A|A|A|A|A|T|T|A|A|C|T|G|A|A|T|T|A|G|A|C|A|T|C|A|A|G|C|A|G|C|C|A|A|G|T|C|A|C|T|C|A|G|A|C|T| 459
QY 418 C|T|G|C|A|G|A|G|A|C|T|T|C|A|C|T|G|C|A|G|C|T|T|C|C|A|G|A|T|G|T|T|C|C|A|G|G|G|A|C|A|T|G|G|C|C|A| 477
Db 460 G|C|C|A|T|G|G|G|A|C|T|C|T|A|C|T|A|C|A|G|T|T|C|C|A|G|A|A|C|C|T|A|C|A|G|G|A|G|A|A|T|G|---T 516
QY 478 G|C|A|G|A|G|A|C|A|G|A|C|A|C|T|G|G|G|A|G|C|T|C|C|G|A|T|A|T|A|A|T|C|T|A|C|A|C|A|A|T|A|T|C|A|C|A|T|G| 537
Db 517 T|C|A|G|A|G|A|C|A|G|A|C|A|C|T|G|T|G|A|C|C|C|T|C|A|T|A|T|A|C|A|T|G|A|A|C|A|A|A|A|T|T|C|C|A|C|A|T|G| 576
QY 538 G|C|C|A|T|G|A|G|T|T|A|C|G|G|A|C|T|C|T|A|G|A|T|T|G|C|C|A|T|G|A|C|T|T|A|C|G|G|A|C|T|T|G|A|G|C|A|C|C|A|T|C| 597
Db 577 G|C|T|G|A|T|G|A|A|T|T|A|-----G|G|A|C|T|C|T|G|A|G|A|A|C|G|A|T|C| 609
QY 598 A|G|A|T|A|T|G|G|C|A|T|T|A|C|A|T|G|A|G|A|G|G|A|T|T|G|T|G|G|G|C|T|G|G|C|T|G|G|C|T|T|A|T|C|T|T|C| 657
Db 610 A|G|A|C|T|G|C|T|A|T|C|C|A|C|A|T|T|G|G|A|G|G|A|G|T|C|T|G|C|T|G|G|G|T|T|G|A|C|C|C|T|G|G|C|A|C|T|T|A|T|C|A|T| 669
QY 658 G|G|G|C|T|T|A|T|T|T|C|A|A|T|G|T|A|T|T|C|T|C|A|T|A|G|C|A|A|G|A|G|A|T|A|C|A|A|T|T|A|A|G|C|T|C| 717
Db 670 G|G|T|G|T|C|T|A|T|C|C|T|A|A|T|G|T|A|T|T|C|T|G|T|A|G|A|A|A|A|A|G|A|G|T|A|T|G|A|G|C|T|T| 729
QY 718 A|T|C|T|T|T|G|C|C|A|C|C|T|C|C|C|C|T|C|A|G|A|T|T|G|C|A|A|T|G|C|A|G|T|A|G|C|A|G|A|G|G|A|T|T|C|G|C| 777
Db 730 A|T|T|A|C|A|C|T|G|G|C|C|A|C|T|G|C|C|C|A|G|G|G|T|G|G|C|A|A|T|G|C|A|G|A|G|C|A|G|T|C|A|G|A|T|T|C|G|C| 789
QY 778 T|C|A|G|A|G|A|A|A|C|A|T|T|A|T|A|C|A|T|T|G|A|G|A|G|A|A|C|G|T|A|T|T|G|A|G|G|A|G|G|A|G|C|C|A|T|G|A|G| 837
Db 790 T|C|T|G|A|G|A|A|A|T|A|T|C|T|A|C|A|C|C|A|T|C|A|G|A|G|A|A|C|G|T|A|T|A|T|G|A|G|A|T|T|C|A|A|A|T|G|A|G| 849
QY 838 T|A|T|A|T|G|C|T|A|T|G|T|C|A|G|C|A|G|C|A|G|C|A|C|C|T|C|A|C|A|C|C|T|T|G|G|G|T|G|T|G|C|T|T|G|C|A| 897
Db 850 T|A|C|T|A|C|T|G|C|T|A|G|T|C|A|C|A|G|C|A|G|C|A|C|C|T|G|A|C|C|G|C|C|T|G|A|C|T|G|C|C|A|C|T|T|T|A| 909
QY 898 A|T|G|C| 901
Db 910 A|G|G|C| 913
```

Search completed: November 21, 2003, 23:34:31  
Job time : 3377.32 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 11:32:52 ; Search time 248.931 Seconds  
(without alignments)  
9792.239 Million cell updates/sec

Title: US-10-004-633-37

Perfect score: 903

Sequence: 1 atgtttccatcatcttcctt.....gttgcgttccttcgcaatgcga 903

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq.19Jun03:\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*

2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*

3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*

4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*

5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*

6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*

7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*

8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*

9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*

10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*

11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*

12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*

13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*

14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*

15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*

16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*

17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*

18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*

19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*

20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*

21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*

22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*

23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*

24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	903	100.0	903	21	AAAS1905
2	903	100.0	903	22	AAI70281
3	903	100.0	903	22	AAF82629
4	903	100.0	903	22	AAAC92155
5	903	100.0	903	22	AAAF23478
6	903	100.0	903	25	ABQ77041
7	903	100.0	1116	25	ABZ68338
8	903	100.0	1756	22	AAAS22714

9	903	100.0	1772	22	AAH15441
10	903	100.0	2236	21	AAAS1904
11	903	100.0	2236	22	AAI70254
12	903	100.0	2236	22	AAF82616
13	903	100.0	2236	22	AAAC90993
14	903	100.0	2236	22	AAF23465
15	903	100.0	2236	22	ABSS5322
16	903	100.0	2236	25	ABQ77040
17	903	100.0	2237	17	AAI38267
18	902.6	100.0	1475	24	ABL98085
19	901.4	99.8	1116	25	ABZ68339
20	901.4	99.8	1800	22	AAAS22478
21	799	88.5	1326	25	AAI33381
22	712	78.8	769	22	AAH04055
23	585	64.8	2091	22	AAK35717
24	522.8	57.9	555	20	AAK35716
25	456	50.5	2710	21	AAAS1898
26	456	50.5	2710	22	AAI70263
27	456	50.5	2710	22	AAF82609
28	456	50.5	2710	22	AAAC90986
29	456	50.5	2710	22	AAF23458
30	456	50.5	2710	24	ABSS5316
31	456	50.5	2710	25	ABQ77037
32	456	50.5	2712	17	AAI38265
33	456	50.5	2725	25	ABZ68328
34	455	50.4	529	20	AAV87937
35	442.6	49.0	843	21	AAAS1918
36	442.6	49.0	843	22	AAI70255
37	442.6	49.0	843	22	AAF82615
38	442.6	49.0	843	22	AAAC90992
39	442.6	49.0	843	22	AAF23464
40	442.6	49.0	843	25	ABQ77038
41	440.8	48.8	862	25	ABZ68329
42	396.6	43.9	1203	21	AAH87799
43	374.2	41.4	438	20	AAAX4096
44	74	8.2	4209	25	AAAC50985
45	72.6	8.0	1079	25	ABZ68335

## ALIGNMENTS

RESULT 1	AAAS1905	standard; cDNA, 903 BP.
ID	AAAS1905	
AC	AAAS1905;	
XX		
DT	31-OCT-2000	(first entry)
DE	Human T helper cell differentially expressed gene 200 cDNA.	
XX		
KW	T helper cell; differential expression; 200 gene; immunomodulator;	
KW	anti-inflammatory; anti-arthritic; antibacterial; immunosuppressive;	
KW	thymomimetic; anti-thyroid; anti-asthmatic; anti-allergic; antiviral;	
KW	protozoacide; lymphocyte; modulator; gene therapy; ss.	
OS	Homo sapiens.	
XX		
XX		
FT	Key	Location/Qualifiers
FT	CDS	1..903
FT		/tag= a
FT		/partial
FT	sig_peptide	1..60
FT		/tag= b
FT	mat_peptide	61..903
FT		/tag= c
XX		
PN	US6084083-A.	
XX		
PD	04-JUL-2000.	
XX		
PF	28-MAR-1997;	97US-0829525.

Human cDNA sequenc  
Human T helper cel  
Human 200 gene, pr  
Human TH1 specific  
Human 200 gene nuc  
Human 200 gene. H  
DNA encoding human  
Human TH1-associat  
200 gene different  
Human polynucleoti  
Nucleotide sequenc  
Human cDNA encodin  
NOVX DNA sequence  
Human cDNA clone (c  
cDNA encoding a pr  
Murine T helper ce  
Murine 200 gene, pr  
Murine TH1 specific  
Murine 200 gene nuc  
Murine 200 gene  
Murine TH1 specific  
Murine TH1-associat  
200 gene different  
Nucleotide sequenc  
EST clone FG708.  
T helper cell diff  
Mouse 200 gene, pr  
Murine TH1 specific  
Murine nucleotide s  
DNA associated wit  
Murine TH1-associat  
Nucleotide sequenc  
Human SECK cDNA C1  
Human secreted pro  
Human bladder canc  
Nucleotide sequenc





PS Disclosure; Column 135-136; 108pp; English.

The present sequence is that of cDNA comprising the complete open reading frame of the human 200 gene. The encoded protein (see *AAIM50223*) is a receptor of the Ig superfamily class. The cDNA was isolated from a human lymphocyte cDNA library using mouse 200 gene cDNA as probe. The nucleotide sequence contained within B. coli clone felt 200C is deposited as ATCC 69967. The human 200 gene is preferentially expressed in mature, fully differentiated T helper 1 subpopulation TH1 cells relative to subpopulation TH2 cells. The gene can be used diagnostically or as a target for therapeutic intervention for the treatment of immune disorders. A claimed method for diagnosing a TH cell subpopulation-related immune disorder involves detecting the level of a 200 gene product, or an RNA encoding it, so that if the level differs from that in a control sample, the disorder is diagnosed. The disorder is especially a TH1 cell subpopulation-related immune disorder, such as multiple sclerosis, psoriasis or insulin-dependent diabetes (claimed). In addition to the 200 gene, the invention provides other genes that are differentially expressed within and among TH cells and TH cell subpopulations and which can be used in methods for the diagnosis, prognosis, evaluation and treatment of TH cell subpopulation-related disorders, for the identification of subjects exhibiting a predisposition to such conditions, for monitoring patients undergoing clinical evaluation for the treatment of such disorders, and for monitoring the efficacy of compounds used in clinical trials. Other immune disorders that can be treated/diagnosed include Crohn's disease, reactive arthritis, Lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis, Grave's disease, contact dermatitis, graft rejection, graft versus host disease, sarcoidosis, atopic conditions, asthma, allergy, allergic rhinitis, food allergy, eosinophilia, conjunctivitis, glomerular nephritis, helminthic infection (e.g. leishmaniasis), viral infection (e.g. HIV), and bacterial infection (e.g. tuberculosis and lepromatous leprosy).

SQ Sequence 903 BP; 242 A; 225 C; 221 G; 215 T; 0 other;

Query Match	Score	DB	Length
100.0%	903	22	903

Matches 903; Conservative 0; Mismatches 0; Indels 0

Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGTTTTCACATTTTCCCTTTGACGTGTCCTGCTGCTGCTACACTTACAAAG	60
Db	1	ATGTTTTCACATTTTCCCTTTGACGTGTCCTGCTGCTGCTACACTTACAAAG	60
Qy	61	TCCTCAGAAAGTGAATACAGACGGAGGTGGTGCAATGCTATCTGCCCCCTTCTAC	120
Db	61	TCCTCAGAAAGTGAATACAGACGGAGGTGGTGCAATGCTATCTGCCCCCTTCTAC	120
Qy	121	ACCCGACGCGCCCCAGGSAACCTCGTCCCGCTCTCTGGGCAAAAGAGCCTGTCCTGTG	180
Db	121	ACCCGACGCGCCCCAGGSAACCTCGTCCCGCTCTCTGGGCAAAAGAGCCTGTCCTGTG	180
Qy	181	TTTGAATGTGGCAACGTGTGCTCAGGACGTGATGAAGGGATGTGAATTTATTTGACATCC	240
Db	181	TTTGAATGTGGCAACGTGTGCTCAGGACGTGATGAAGGGATGTGAATTTATTTGACATCC	240
Qy	241	AGATACGTGGCTAAATGCGGATTTCCGCCAAGGAGATGTGTCTCCCTGACCATAGAGATGTG	300
Db	241	AGATACGTGGCTAAATGCGGATTTCCGCCAAGGAGATGTGTCTCCCTGACCATAGAGATGTG	300
Qy	301	ACCTTAGCAGACGATGGGATCTACTGCTGCGGATCTCAATCCAGGACATAATGAT	360
Db	301	ACCTTAGCAGACGATGGGATCTACTGCTGCGGATCTCAATCCAGGATATATGAT	360
Qy	361	GAAAAATTTTAACTGAAAGTTGTGATCAAAACGACCAAGTCAACCCCTGCACCGACTCTG	420
Db	361	GAAAAATTTTAACTGAAAGTTGTGATCAAAACGACCAAGTCAACCCCTGCACCGACTCTG	420
Qy	421	CAGAGACATTCACCTGACGCTTTCCAAAGATGCTTACCAACAGGGGACATGCGCCACGA	480
Db	421	CAGAGACATTCACCTGACGCTTTCCAAAGATGCTTACCAACAGGGGACATGCGCCACGA	480

OY	481	GAGCACACGACACTGGGAGGCGTCCCTGATATTAATCTAACCAATATCCACATGGCC	540
Db	481	GAGCACACGACACTGGGAGGCGTCCCTGATATTAATCTAACCAATATCCACATGGCC	540
OY	541	AATGAGTTACGGGACTCTTAGATTGGCCAAATGACTTACGGGACTCTGGAGCAACCATCAG	600
Db	541	AATGAGTTACGGGACTCTTAGATTGGCCAAATGACTTACGGGACTCTGGAGCAACCATCAG	600
OY	601	ATAGGCATCTACATCGAGACAGGAGATCTGTGGTGGGCTGGCTCTTATCTTCGGG	660
Db	601	ATAGGCATCTACATCGAGACAGGAGATCTGTGGTGGGCTGGCTCTTATCTTCGGG	660
OY	661	GCTTTAATTTTCAATGGTATTTCTCATGCAAGAAGAAATACAGAAITTTAAAGCCTCATC	720
Db	661	GCTTTAATTTTCAATGGTATTTCTCATGCAAGAAGAAATACAGAAITTTAAAGCCTCATC	720
OY	721	TCTTTGGCCAACTCCCTCCCTCAGAGATTGGCAATGCAAGTGCAGAGGGAATTCGCTCA	780
Db	721	TCTTTGGCCAACTCCCTCCCTCAGAGATTGGCAATGCAAGTGCAGAGGGAATTCGCTCA	780
OY	781	GAAGAAAACATCTATACCATTTGAAGAAGACGTATATGAAGTGGAGAGCCCAATGAGTAT	840
Db	781	GAAGAAAACATCTATACCATTTGAAGAAGACGTATATGAAGTGGAGAGCCCAATGAGTAT	840
OY	841	TATTGCTATGTGACGACGACGAGCAAGCAACCTTCACACCTTTGGGTTGTGCTTTGCAATG	900
Db	841	TATTGCTATGTGACGACGACGAGCAAGCAACCTTCACACCTTTGGGTTGTGCTTTGCAATG	900
OY	901	CCA 903	
Db	901	CCA 903	

### RESULT 3

ID AAF82629 standard; cDNA; 903 BP

AC AAF82629

DT 18-JUN-2001 (first entry)

DE Human TH1 specific 200 gene open reading frame

KM Human; T helper cell; TH cell; TH1; TH2; immunomodulator;  
KM anti-inflammatory; anti-alleergic; dermatological; antiviral;  
KM antibacterial; T helper lymphocyte modulator; gene therapy;  
KM TH specific gene; 200 gene; immune disorder; inflammation;  
KM infection; ss.

OS Homo sapiens.

FH	Key	Location/Qualifiers
22		

/\*tag= a

[illegible]

XX 20 MAR 2001

XX 01 MAR 1960

XX  
03 MAR 1968  
030833

PR 07-JUN-1995; 95US-0487748.  
YY

PA (MILL-) MILLENNIUM PHARM INC.  
XX

PI Levinson DA;  
xy

DR WPI; 2001-272703/28  
DR P-PSDB: AABR1518

XX

QY 1 ATGTTTCACATCTTCCCTTTGACTGTGTCTCTGCTGCTGCTACTACTTACAAG 60





Oy		301	CTCTAGACGACAGTGGGATCTACTGCTGCCGGATTCCAAATCCCAAGGCATAATGATGTAT	360
Db		301	ACTTAGCAGACAGTGGGATCTACTGCTGCCGGATCCAATCCCAAGGCATAAAGAATGAT	360
Oy		361	GAATAATTTAACCTTAAGTTGGTGATCAAACCAGCGAAGSTGACCCCTGCACCGACTCTG	420
Db		361	GAAAAATTTAACCTTAAGTTGGTGATCAACAACGCGAAGSTGACCCCTGCACCGACTCTG	420
Oy		421	CAGAGAGACTTCACCTGCAGACCTTTCCMAGATGCTTTACCAACGAGGGAGCATGGCCCCAGA	480
Db		421	CAGAGAGACTTCACCTGCAGACCTTTCCMAGATGCTTTACCAACGAGGGAGCATGGCCCCAGA	480
Oy		481	GAGACACAGACACTGGGGAGCCTCCTCGATATATTAATCTTAACCAAAATATTCACATTGGCC	540
Db		481	GAGACACAGACACTGGGGAGCCTCCTCGATATATTAATCTTAACCAAAATATTCACATTGGCC	540
Oy		541	AATGAGTTACGGGACTCTAGATTGGCCATATGACTTACGGGGACTCTGAGCAACCATCAGA	600
Db		541	AATGAGTTACGGGACTCTAGATTGGCCATATGACTTACGGGGACTCTGAGCAACCATCAGA	600
Oy		601	ATAGGCATCTACATCGAGCAGAGGAGATCTGTGCTGGCTGGCTCTGCTCTTAATCTTCGGC	660
Db		601	ATAGGCATCTACATCGAGCAGAGGAGATCTGTGCTGGCTGGCTCTGCTCTTAATCTTCGGC	660
Oy		661	GCTTTAATTTTCAAAATGSTATTTCTCATAGCAAAAAGAGAGATACAGAAATTTAAGCTTCATC	720
Db		661	GCTTTAATTTTCAAAATGSTATTTCTCATAGCAAAAAGAGATACAGAAATTTAAGCTTCATC	720
Oy		721	TCTTTGGCCAACTCCCTCCCTCAGATTTGGCAAAATGCAAGTATGCAAGGAATTTCCCTCA	780
Db		721	TCTTTGGCCAACTCCCTCCCTCAGATTTGGCAAAATGCAAGTATGCAAGGAATTTCCCTCA	780
Oy		781	GAAGAAACATCTATACCATTTAACAAGAAACGTATATGAAGTGGAGAGCCCAATGAGTAT	840
Db		781	GAAGAAACATCTATACCATTTAACAAGAAACGTATATGAAGTGGAGAGCCCAATGAGTAT	840
Oy		841	TATTGCTATGTACGACGACGAGCAACCCCTCAACACTTTGGGTTGTGCTTTGCATG	900
Db		841	TATTGCTATGTACGACGACGAGCAACCCCTCAACACTTTGGGTTGTGCTTTGCATG	900
Oy		901	CCA 903	
Db		901	CCA 903	
<hr/>				
RESULT 7				
ABZ68338	ID	ABZ68338	standard; DNA; 1116 BP.	
XX	XX	ABZ68338;		
XX	AC	22-APR-2003	(first entry)	
XX	DE	Nucleotide sequence of human TIM-3 allele 1.		
XX	XX	T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;		
KW	XX	myeloid/plastic syndrome; chromosome 5; malignancy; allergy; eczema;		
KM	XX	allergic T cell response; autoimmune disease; gene; ss.		
OS	XX	Homo sapiens.		
FH	XX	Key	Location/Qualifiers	
FT	CDS	58..963	/tag= "a"	
FT		/product= "TIM-3"		
FT		/transl_except= (pos:475..477,aa:Arg)		
PN	XX	WO2003002722-A2.		
PD	XX	09-JAN-2003.		
PF	XX	01-JUL-2002; 2002WO-US20890.		

XX 29-JUN-2001; 2001US-302344P.  
FR  
XX  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
XX  
PI McIntire JI, Detryff RH, Unetsu DT, Freeman GJ, Kuchoo V;  
DR WPI; 2003-210268/20.  
XX P-PSDB; ABP70444.  
XX  
PT New nucleic acid comprising a mammalian T cell Immunoglobulin domain  
PT and Mucin domain gene sequences, useful for treating cancer or asthma,  
PT allergy, eczema or autoimmune disease -  
XX  
PS  
XX Claim 5; Page 89-90; 94pp; English.

CC The present sequence encodes a human T cell immunoglobulin domain and  
CC mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1,  
CC TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with  
CC conserved IGV and mucin domains. The locus comprising the TIM family is  
CC genetically associated with immune dysfunction, including asthma. The  
CC TIM gene family is located within a region of human chromosome 5 that  
CC is commonly deleted in malignancies and myelodysplastic syndrome.  
CC Variants of TIM-1 and TIM-3 are associated with susceptibility to  
CC atway hyperreactivity and allergic T cell responses, and other variants  
CC associated with protection against these responses. T cells express TIM  
CC proteins, which critically regulate CD4 T cell differentiation. Th1  
CC cells preferentially express TIM-3, while Th2 cells preferentially  
CC express TIM-1. TIM polypeptides and polynucleotides are useful for  
CC treating cancer, asthma, allergies, eczema or autoimmune diseases.  
XX

SQ Sequence 116 BP; 291 A; 268 C; 270 G; 287 T; 0 other;

Query Match	100.0%	Score 903	DB 25	Length 1116
Best Local Similarity	100.0%	Pred. No. 5.3e-276		
Matches	903	Conservative	0	Mismatches 0; Indels 0; Gaps 0
QY	1	ATGTTTTCACACATCTCCCTTGACTGTGTCCTGCTGCTCTGCTGCTACTACTTACCAAG	60	
Db	58	ATGTTTTCACATCTTCCTTTGACTGTGTCTCTGCTGCTGCTGCTGCTACTACTTACCAAG	117	
QY	61	TCCTCAGAAGTGGAATACAGAGCGAGGTGGTTCAGAAATGCTATCTGCTTCCTCTAC	120	
Db	118	TCCTCAGAAGTGGAATACAGAGCGGAGGTGGTTCAGAAATGCTATCTGCTTCCTCTAC	177	
QY	121	ACCCGAGCGCCCGAGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	180	
Db	178	ACCCGAGCGCCCGAGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	237	
QY	181	TTTGAATGTGGCACTGTGTGCTCAGGACTGATGAAAGGAGTGTGAATTATTTGACATCC	240	
Db	238	TTTGATGTGGCAAGTGTGTGCTCAGGACTGATGAAAGGAGTGTGAATTATTTGACATCC	297	
QY	241	AGATCTGGCTAAATGAGGGATTTCCGCAAAAGAGATGTGTCCTGACCAATAGGAATGTG	300	
Db	298	AGATCTGGCTAAATGAGGGATTTCCGCAAAAGAGATGTGTCCTGACCAATAGGAATGTG	357	
QY	301	ACTTAGCAGACAGTGGGATCTACTGCTGCGGATCCAAATCCACAGCATTAATGAATGAT	360	
Db	358	ACTTAGCAGACAGTGGGATCTACTGCTGCGGATCCAAATCCACAGCATTAATGAATGAT	417	
QY	361	GAATAATTTAACTGAAGTTGTCATCAAAACGACCAAGGTCACTCCCTGACCGGACTTG	420	
Db	418	GAATAATTTAACTGAAGTTGTCATCAAAACGACCAAGGTCACTCCCTGACCGGACTTG	477	
QY	421	CAGAGAGCTTACCTGACGCTTTTCCAGAGATGTTTACACACAGGGGACATGGCCACGCA	480	
Db	478	CAGAGAGCTTACCTGACGCTTTTCCAGAGATGTTTACACACAGGGGACATGGCCACGCA	537	
QY	481	GAGACACAGACACTGGGGAGCTCCCTGTATTAATCTAAACAAATATCAATTGGCC	540	
Db	538	GAGACACAGACACTGGGGAGCTCCCTGTATTAATCTAAACAAATATCAATTGGCC	597	



















CC present sequence represents the DNA sequence encoding the human 200  
CC gene protein of the invention.

Search completed: November 21, 2003, 19:49:12

**SQ** Sequence 2236 BP; 586 A; 518 C; 525 G; 606 T; 1 other;

Query Match	100.0%	Score 903;	DB 24	Length 2236;
Best Local Similarity	100.0%	Pred. NC 7.7e-276;		
Matches 903;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY		1	ATGTTTTCACATCTTCCCTTTGACTGTGCCTGGCTGCTGCTGCTACTAATTACAAG	60
Db		42	ATGTTTTCACATCTTCCCTTTGACTGTGCCTGGCTGCTGCTGCTACTAATTACAAG	101
QY		61	TCTCTGAAGTGGAATACAGACCGAAGTCCGTCAAAATGCCATCTTCCCTGCTTAC	120
Db		102	TCTCTGAAGTGGAATACAGACCGAAGTCCGTCAAAATGCCATCTTCCCTGCTTAC	161
QY		121	ACCCCAAGCCCCCAGAGGAACCTCGTGCCTGTCTGCTGGGCAAGAACCCTGTCTGTG	180
Db		162	ACCCCAAGCCCCCAGAGGAACCTCGTGCCTGTCTGCTGGGCAAGAACCCTGTCTGTG	221
QY		181	TTTGAATGTGGCAACGTGGTGCTCAAGCATGTATGAAGAAGATTTAATGGACATCC	240
Db		222	TTTGAATGTGGCAACGTGGTGCTCAAGCATGTATGAAGAAGATTTAATGGACATCC	281
QY		241	AGATACTGGCTAAATGGGGATTTTCGCAAAAGAGATGTGTCCCTGACCATAGAAATGTG	300
Db		282	AGATACTGGCTAAATGGGGATTTTCGCAAAAGAGATGTGTCCCTGACCATAGAAATGTG	341
QY		301	ACTCTTGCAAGACAGTGGGATCTTAAGTCTGCTCGGGAATCCAATCCAGGCAATATGAT	360
Db		342	ACTCTTGCAAGACAGTGGGATCTTAAGTCTGCTCGGGAATCCAATCCAGGCAATATGAT	401
QY		361	GAAAAATTTAACCTGAAGTTGGTCAATCAAACAGCAAGTCAACCTCGACCGAAGCTGTG	420
Db		402	GAAAAATTTAACCTGAAGTTGGTCAATCAAACAGCAAGTCAACCTCGACCGAAGCTGTG	461
QY		421	CAGAGAGACTTCACTGCAAGCTTTTCCAAAGATGCTTACACACAGGGGACATGGCCAGCA	480
Db		462	CAGAGAGACTTCACTGCAAGCTTTTCCAAAGATGCTTACACACAGGGGACATGGCCAGCA	521
QY		481	GAGACACAGACACTGGGGAGCCTCCCTGATATAAATCTTAACAATAATCCATTGGCC	540
Db		522	GAGACACAGACACTGGGGAGCCTCCCTGATATAAATCTTAACAATAATCCATTGGCC	581
QY		541	AATGAAATTACGGGACTCTGATTTGGCCAAATGACTTAACGGGACTCTGAGCAACCATAGA	600
Db		582	AATGAAATTACGGGACTCTGATTTGGCCAAATGACTTAACGGGACTCTGAGCAACCATAGA	641
QY		601	ATAGGCAATCAATCGAGACAGAGATCTGTGCTGGGCTGGCTGTGCTTATCTTGAGC	660
Db		642	ATAGGCAATCAATCGAGACAGAGATCTGTGCTGGGCTGGCTGTGCTTATCTTGAGC	701
QY		661	GCTTTAATTTCAAAATGATTTCTATAGCAAAAGAAAGATACAAATTTAAGCCTCATC	720
Db		702	GCTTTAATTTCAAAATGATTTCTATAGCAAAAGAAAGATACAAATTTAAGCCTCATC	761
QY		721	TCTTTGGCCAACCTCCCTCCCTCAGAGATTTGGCAATGCAATGACAGAGGAATTCGCTCA	780
Db		762	TCTTTGGCCAACCTCCCTCCCTCAGAGATTTGGCAATGCAATGACAGAGGAATTCGCTCA	821
QY		781	GAGAAAACATCTATACATTTGAGGAAACGPATATGAAGTGAAGAGGCCAATGAGTAT	840
Db		822	GAGAAAACATCTATACATTTGAGGAAACGPATATGAAGTGAAGAGGCCAATGAGTAT	881
QY		841	TATGTCTATGTGACGACAGGAGCAACCTTCACAACCTTTGGGTGTGTGCTTTGCAATG	900
Db		882	TATGTCTATGTGACGACAGGAGCAACCTTCACAACCTTTGGGTGTGTGCTTTGCAATG	941
QY		901	CCA 903	
Db		942	CCA 944	

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 16:48:30 ; Search time 55.4846 Seconds  
(without alignments)  
7183.411 Million cell updates/sec

Title: US-10-004-633-37  
Perfect score: 903

Sequence: 1 atgtttcacattccctt.....gttgctgcttgcacatgccca 903

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	903	100.0	903	US-08-829-525-37	Sequence 37, Appl
2	903	100.0	903	US-08-609-583A-37	Sequence 37, Appl
3	903	100.0	903	US-08-937-399-37	Sequence 37, Appl
4	903	100.0	903	US-09-310-367-37	Sequence 37, Appl
5	903	100.0	903	US-09-033-337-37	Sequence 37, Appl
6	903	100.0	903	US-09-464-231-37	Sequence 37, Appl
7	903	100.0	2236	US-08-829-525-23	Sequence 23, Appl
8	903	100.0	2236	US-08-609-583A-23	Sequence 23, Appl
9	903	100.0	2236	US-08-937-399-23	Sequence 23, Appl
10	903	100.0	2236	US-09-310-367-23	Sequence 23, Appl
11	903	100.0	2236	US-09-033-337-23	Sequence 23, Appl
12	903	100.0	2236	US-09-464-231-23	Sequence 23, Appl
13	456	50.5	2710	US-08-487-748A-8	Sequence 8, Appl
14	456	50.5	2710	US-08-480-070C-8	Sequence 8, Appl
15	456	50.5	2710	US-08-829-525-8	Sequence 8, Appl
16	456	50.5	2710	US-08-609-583A-8	Sequence 8, Appl
17	456	50.5	2710	US-08-937-399-8	Sequence 8, Appl
18	456	50.5	2710	US-09-310-367-8	Sequence 8, Appl
19	456	50.5	2710	US-09-033-337-8	Sequence 8, Appl
20	456	50.5	2710	US-09-464-231-8	Sequence 8, Appl
21	442.6	49.0	843	US-08-829-525-22	Sequence 22, Appl
22	442.6	49.0	843	US-08-609-583A-22	Sequence 22, Appl
23	442.6	49.0	843	US-08-937-399-22	Sequence 22, Appl
24	442.6	49.0	843	US-09-310-367-22	Sequence 22, Appl
25	442.6	49.0	843	US-09-033-337-22	Sequence 22, Appl
26	442.6	49.0	843	US-09-464-231-22	Sequence 22, Appl
27	64.4	7.1	1380	US-09-620-312D-59	Sequence 59, Appl

28	56.4	6.2	2093	1	US-08-287-001A-1	Sequence 1, Appl
29	56.4	6.2	2093	3	PCT-US95-09941-1	Sequence 1, Appl
30	37	4.1	578	5	US-09-085-199B-25	Sequence 25, Appl
31	35.2	3.9	2091	1	US-08-765-081-4	Sequence 4, Appl
32	35.2	3.9	2091	3	US-09-098-082-4	Sequence 4, Appl
33	35.2	3.9	2091	5	PCT-US95-06994-4	Sequence 4, Appl
34	35.2	3.9	8041	1	US-08-765-081-1	Sequence 1, Appl
35	35.2	3.9	8041	3	US-09-098-082-1	Sequence 1, Appl
36	35.2	3.9	8041	5	PCT-US95-06994-1	Sequence 1, Appl
37	35.2	3.9	87563	4	US-09-453-702B-57	Sequence 57, Appl
38	34.4	3.8	3502	2	US-08-724-394A-16	Sequence 16, Appl
39	32.6	3.6	1664976	4	US-08-916-421B-1	Sequence 1, Appl
40	31.8	3.5	2175	4	US-08-914-999-3	Sequence 3, Appl
41	31.8	3.5	90050	3	US-09-245-041-5	Sequence 5, Appl
42	31.4	3.5	2310	1	US-08-461-680B-1	Sequence 1, Appl
43	31.4	3.5	2311	1	US-08-103-445-1	Sequence 1, Appl
44	31.4	3.5	3257	5	PCT-US91-09784-1	Sequence 1, Appl
45	31.2	3.5	418	3	US-09-141-000-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1  
US-08-829-525-37  
Sequence 37, Application US/08829525  
Patent No. 6084083  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/829, 525  
FILING DATE: 28-MAR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/609, 583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487, 748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398, 633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-081  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 903 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-829-525-37  
Query Match 100.0%; Score 903; DB 3; Length 903;





```

OY 421 CAGAGAGACTTCACCTGACACCTTTCCAAAGATGCTTACACACGAGGGGAACTGGCCAGCA 480
Db 421 CAGAGAGACTTCACCTGACACCTTTCCAAAGATGCTTACACACGAGGGGAACTGGCCAGCA 480
OY 481 GAGACACAGACA CTGGGGAGCCTCCCTGATATATTAATCTAACACAAATATTCACATTGGCC 540
Db 481 GAGACACAGACA CTGGGGAGCCTCCCTGATATATTAATCTAACACAAATATTCACATTGGCC 540
OY 541 AATGAGTTACGGGACTCTTAGATTGGCCATGACTTACGGGACTCTGGAGCAACCATCAGA 600
Db 541 AATGAGTTACGGGACTCTTAGATTGGCCATGACTTACGGGACTCTGGAGCAACCATCAGA 600
OY 601 ATAGGACATCTACATCGGAGCAGGAGATCTGTGCTGGGCTGGCTGGCTCTTATCTTCGGC 660
Db 601 ATAGGACATCTACATCGGAGCAGGAGATCTGTGCTGGGCTGGCTGGCTCTTATCTTCGGC 660
OY 661 GCTTAAATTTCAAAATGATGATTTCTCATAGCAAGAGAGATACAGAAATTTAAGCCTCATC 720
Db 661 GCTTAAATTTCAAAATGATGATTTCTCATAGCAAGAGATACAGAAATTTAAGCCTCATC 720
OY 721 TCTTTGGCCACCTCCCTCCCTTCAGATTTGGCAATGCAATGCAATGCAAGGAAATTTGCTCA 780
Db 721 TCTTTGGCCACCTCCCTCCCTTCAGATTTGGCAATGCAATGCAATGCAAGGAAATTTGCTCA 780
OY 781 GAAAGAAAACATCATATATCAATTTGAAAGAAAGTATATGAAGTGAAGGAGCCCAATGAGTAT 840
Db 781 GAAAGAAAACATCATATATCAATTTGAAAGAAAGTATATGAAGTGAAGGAGCCCAATGAGTAT 840
OY 841 TATTCCTATGTCAGCAGCAGGAGCAACCTTCACAACTTTGGGTTGTGCTTTGCATG 900
Db 841 TATTCCTATGTCAGCAGCAGGAGCAACCTTCACAACTTTGGGTTGTGCTTTGCATG 900
OY 901 CCA 903
Db 901 CCA 903

RESULT 3
US-08-937-399-37
; Sequence 37, Application US/08937399
; Patent No. 6288218
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITILE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fasteq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,399
; FILING DATE:
; CLASSIFICATION:
; APPLICATION DATA:
; APPLICATION NUMBER: 08/609,583
; FILING DATE: 01-MAR-1996
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742

```

	/	REFERENCE/DOCKET NUMBER:	7853-048	
	/	TELECOMMUNICATION INFORMATION:		
	/	TELEPHONE:	212-790-9090	
	/	TELEX:	212-869-8864	
	/	TELEX:	66141 PENNIE	
	/	INFORMATION FOR SEQ ID NO:	37:	
	/	SEQUENCE CHARACTERISTICS:		
	/	LENGTH:	903 base pairs	
	/	TYPE:	nucleic acid	
	/	STRANDEDNESS:	singlne	
	/	TOPOLOGY:	linear	
	/	MOLECULE TYPE:	DNA	
	/	US-08-937-399-37		
		Query Match	100.0%; Score 903;	DB 3; Length 903;
		Best Local Similarity	100.0%; Pred. No.	1.3e+297;
		Matches 903; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
OY		1	AAGTTTACACTCTCCCTTGATCGTCGTCGTGCCTCACTAACAAG	60
Db		1	AAGTTTTACACTCTCCCTTTGACTGGTCTCCTGCTGCTGCATCTAAAGA	60
OY		61	TCTTCAGAAAGGAATAACAGACCGGAGTCGGTCAGATTGCCATTCTCCCTGTTAC	120
Db		61	TCTTCAGAAAGTGAATAACAGACCGGAGTCGGTCAGATTGCCATTCTCCCTGTTAC	120
OY		121	ACCACGACCGCCCCAGGGAACTCGTGCCGTCCTGTCGGGCAAAGAACCTGTCGTG	180
Db		121	ACCACGACCGCCCCAGGGAACTCGTGCCGTCCTGTCGGGCAAAGAACCTGTCGTG	180
OY		181	TTTGAATGTGGCAACGTGTGCTCAGACCTGATGAAGGATGTGTAATTTTGAATCC	240
Db		181	TTTGAATGTGGCAACGTGTGCTCAGACCTGATGAAGGATGTGTAATTTTGAATCC	240
OY		241	AGATCTGGCTAAATAGGGGATTTCCGCAAGAGAGATGTGTCCTGACCATAGAGATG	300
Db		241	AGATCTGGCTAAATAGGGGATTTCCGCAAGAGAGATGTGTCCTGACCATAGAGATG	300
OY		301	ACTTAGCACAGACGTGGGATCTACTGCTCGCCGATCCAAATCCCAGCATATGATAT	360
Db		301	ACTTAGCACAGACGTGGGATCTACTGCTCGCCGATCCAAATCCCAGCATATGATAT	360
OY		361	GAAAAATTAAACCTGAATTTGTCATCAACACGCAAGGTCACCCCTGCACCGACTCG	420
Db		361	GAAAAATTAAACCTGAATTTGTCATCAACACGCAAGGTCACCCCTGCACCGACTCG	420
OY		421	CAGAGAGCTTCACTGCAAGCTTTCCAAAGANTCTTACACACAGGGGCAATGGCCACA	480
Db		421	CAGAGAGCTTCACTGCAAGCTTTCCAAAGANTCTTACACACAGGGGCAATGGCCACA	480
OY		481	GAGACACAGACACTGGGGAGCTCCCTCATATTAATCTTAACAAATATCCATTGGCC	540
Db		481	GAGACACAGACACTGGGGAGCTCCCTCATATTAATCTTAACAAATATCCATTGGCC	540
OY		541	AATGAGTTAGCGGACTTAGATTGGCCAATGACTTACCGGACTCTGAGCAACATCAGA	600
Db		541	AATGAGTTAGCGGACTTAGATTGGCCAATGACTTACCGGACTCTGAGCAACATCAGA	600
OY		601	ATAGGACATCTACNTCGAGACAGGATCTGTGCTGGGCTCGGCTCTTATCTTGGGC	660
Db		601	ATAGGACATCTACNTCGAGACAGGATCTGTGCTGGGCTCGGCTCTTATCTTGGGC	660
OY		661	GCTTAAATTTTCAATGSTATTTCTCATAGCAAGAGATACAGAAATTTAAGCTCATC	720
Db		661	GCTTAAATTTTCAATGSTATTTCTCATAGCAAGAGATACAGAAATTTAAGCTCATC	720
OY		721	TCTTTGGCAACTTCCTCCCTCAGAGATTGGCAATGCAATGCAAGAGAGGAAATTCGCTCA	780
Db		721	TCTTTGGCAACTTCCTCCCTCAGAGATTGGCAATGCAATGCAAGAGAGGAAATTCGCTCA	780
OY		781	GAAAGAAAATCTATACATTGAAGAGAAAGTATTAAGTGGAGAGACCCCAATGATAT	840
Db		781	GAAAGAAAATCTATACATTGAAGAGAAAGTATTAAGTGGAGAGACCCCAATGATAT	840







Db	222	TTTGAATGCGCAACGTGTGTGCTCAGGACTGATGAAAGGATGTGAATTATTTGGACATCC	281
Oy	241	AGATACCTGCTTAAATBGGGATTTCCGCAAAAGAGATGTGTCCCTGACCATAGAAATGTG	300
Db	282	AGATACCTGCTTAAATBGGGATTTCCGCAAAAGGAGATGTGTCCCTGACCATAGAAATGTG	341
Oy	301	ACTTAGCAGACAGTGGGATCTTACTGCTGCCGGAATCCAAATCCCAAGCANTAAAGAAATGAT	360
Db	342	ACTTAGCAGACAGTGGGATCTTACTGCTGCCGGAATCCAAATCCCAAGCANTAAAGAAATGAT	401
Oy	361	GAATAATTTAAACCTGAAGTGTGTCATCAAAACCGCAAGGTCAACCCCTGCACCGCATCTG	420
Db	402	GAATAATTTAAACCTGAAGTGTGTCATCAAAACCGCAAGGTCAACCCCTGCACCGCATCTG	461
Oy	421	CAGAGAGACTTCACCTGCACGACCTTTTCCAGAGATGCTTACCAACGAGGGACATGGCCGACA	480
Db	462	CAGAGAGACTTCACCTGCACGACCTTTTCCAGAGATGCTTACCAACGAGGGACATGGCCGACA	521
Oy	481	GAGACACAGACACCTGGGGAGCCTCCCTGATATTAATCTTAAACAAATATTCACATTTGGCC	540
Db	522	GAGACACAGACACCTGGGGAGCCTCCCTGATATTAATCTTAAACAAATATTCACATTTGGCC	581
Oy	541	AATGAGTTACGGGACCTCTAGATTGGCCANTGATTTACGGGACCTGTGAGCAACATTCAGA	600
Db	582	AATGAGTTACGGGACCTCTAGATTGGCCANTGATTTACGGGACCTGTGAGCAACATTCAGA	641
Oy	601	ATAGGCATCTACATCGGACAGGAGATCTGTGCTGGGCTGGCTCTGCTCTTATCTTGGGC	660
Db	642	ATAGGCATCTACATCGGACAGGAGATCTGTGCTGGGCTGGCTCTGCTCTTATCTTGGGC	701
Oy	661	GCTTTAATTTTCAAAATGTAATCTCTCATPACAAAGAGAGATACAGAAATTTAAGCCTTCATC	720
Db	702	GCTTTAATTTTCAAAATGTAATCTCTCATPACAAAGAGAGATACAGAAATTTAAGCCTTCATC	761
Oy	721	TCTTTGGCCCAACCTCCCTGCTCAGGATTTGGCAAAAGCAGTACAGAGGGAAATTCGCTCA	780
Db	762	TCTTTGGCCCAACCTCCCTGCTCAGGATTTGGCAAAAGCAGTACAGAGGGAAATTCGCTCA	821
Oy	781	GAAGAAAAATCATCTATACATTGAGAGAAAGATATATGAGTGGAGAGCCCAATGAGTAT	840
Db	822	GAAGAAAAATCATCTATACATTGAGAGAAAGATATATGAGTGGAGAGCCCAATGAGTAT	881
Oy	841	TATTGCTATGTACGACGACGAGCAACCTTCAAACTTTGGGTTGTGCTTTGCAATG	900
Db	882	TATTGCTATGTACGACGACGAGCAACCTTCAAACTTTGGGTTGTGCTTTGCAATG	941
Oy	901	CCA 903	
Db	942	CCA 944	

/  
 / RESULT 8  
 / US-08-609-583A-23  
 / Sequence 23, Application US/08609583A  
 / Patent No. 6204371  
 / GENERAL INFORMATION:  
 / APPLICANT: Levinson, Douglas A.  
 / TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 / TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
 / NUMBER OF SEQUENCES: 37  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Pennie & Edmonds  
 / STREET: 1155 Avenue of the Americas  
 / CITY: New York  
 / STATE: New York  
 / COUNTRY: USA  
 / ZIP: 10036/2711  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Diskette  
 / COMPUTER: IBM Compatible  
 / OPERATING SYSTEM: DOS  
 / SOFTWARE: FastSeq Version 2.0  
 / CURRENT APPLICATION DATA:  
 /

```

1      APPLICATION NUMBER: US/08/609,583A
2      FILING DATE: 01-MAR-1996
3      CLASSIFICATION: 536
4      PRIOR APPLICATION DATA:
5      APPLICATION NUMBER: US 08/487,748
6      FILING DATE: 07-JUN-1995
7      APPLICATION NUMBER: US 08/398,633
8      FILING DATE: 03-MAR-1995
9      ATTORNEY/AGENT INFORMATION:
10     NAME: Cortuzzi, Laura A.
11     REGISTRATION NUMBER: 30,742
12     REFERENCE/DOCKET NUMBER: 7853-048
13     TELECOMMUNICATION INFORMATION:
14     TELEPHONE: 212-790-9090
15     TELEFAX: 212-869-8864
16     TELEX: 66141 PENNIE
17     INFORMATION FOR SEQ ID NO: 23:
18     SEQUENCE CHARACTERISTICS:
19     LENGTH: 2236 base pairs
20     TYPE: nucleic acid
21     STRANDEDNESS: single
22     TOPOLOGY: linear
23     MOLECULE TYPE: DNA
24     FEATURE:
25     NAME/KEY: Coding Sequence
26     LOCATION: 42...944
27     OTHER INFORMATION: Human 200 gene nucleotide
28     OTHER INFORMATION: sequence
29     US-08-609-583A-23

```

Query Match	100.0%	Score 903	DB 3	Length 2236
Best Local Similarity	100.0%	Pred. No. 2.3e-297		
Matches 903	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ATGTTTTCACATCTTCTCCCTTGA	CTGTGCTCTGTGCTGTCTACTACTTACAAAG	60
DB	42	ATGTTTTCACATCTTCTCCCTTGA	CTGTGCTCTGTGCTGTCTACTACTTACAAAG	101
QY	61	TTCTTCAGAACTGTGAATACAGACG	GAAGTCCGTTCAGAAATGCTTATCTG	120
DB	102	TTCTTCAGAACTGTGAATACAGACG	GAAGTCCGTTCAGAAATGCTTATCTG	161
QY	121	ACCCGAGCCGCGCCAGGGAACTT	CGTGCCTGTCTGCGGGCAAGAGACCTG	180
DB	162	ACCCGAGCCGCGCCAGGGAACTT	CGTGCCTGTCTGCGGGCAAGAGACCTG	221
QY	181	TTTGAATGTGGCAACGTGTGCTC	AGACTGATGAAGGATGTGAATTTATG	240
DB	222	TTTGAATGTGGCAACGTGTGCTC	AGACTGATGAAGGATGTGAATTTATG	281
QY	241	AGATATCTGGCTTAAATGGGGAT	TTTCCGCAAGAGATGTGTCCCTGAC	300
DB	282	AGATATCTGGCTTAAATGGGGAT	TTTCCGCAAGAGATGTGTCCCTGAC	341
QY	301	ACTCTAGCAGACAGTGGGATCTA	CTGCTGCGCGGATTCCAATCCGACG	360
DB	342	ACTCTAGCAGACAGTGGGATCTA	CTGCTGCGCGGATTCCAATCCGACG	401
QY	361	GAATAATTTAACTGAAGTTGTC	ATCATCAACACAGCCAAAGTCA	420
DB	402	GAATAATTTAACTGAAGTTGTC	ATCATCATCAACACAGCCAAAGTCA	461
QY	421	CAGAGAGCTTCACTGCAGCCTT	TCCAAAGATGCTTACACACAGGGG	480
DB	462	CAGAGAGCTTCACTGCAGCCTT	TCCAAAGATGCTTACACACAGGGG	521
QY	481	GAGACACAGACACTGGGGAGCT	CCCTCGATTTAATCTAACAAATAT	540
DB	522	GAGACACAGACACTGGGGAGCT	CCCTCGATTTAATCTAACAAATAT	581
QY	541	AATGAGTTACGGGACTCTTAAT	TGCGCAATGACTTTACGGGACTTG	600
DB	582	AATGAGTTACGGGACTCTTAAT	TGCGCAATGACTTTACGGGACTTG	641

Qy	601	ATTGAGCATCTACATCCGAGCAGGAGATCTGCTGGGCTGAGCTGAGCTCTTATCTTGAGC	668
Db	642	ATAGGATCTACATCCGAGCAGGAGATCTGCTGGGCTGAGCTGAGCTCTTATCTTGAGC	701
Qy	661	GCTTTAATTTTCAATGATGATTTCTCATAGCAAAAGAGATACAGATTTAAGCTCATC	720
Db	702	GCTTTAATTTTCAAAAGTATTTCTCATAGCAAAAGAGATACAGATTTAAGCTCATC	761
Qy	721	TCTTTGGCCAACCTCCCTCCCTCAGGATTGGCAATGACATGACAGAGGAAATTCGCTCA	780
Db	762	TCCTTTGGCCAACCTCCCTCCCTCAGGATTGGCAATGACATGACAGAGGAAATTCGCTCA	821
Qy	841	TATTGCTATGTCAGCAGCAGGAGCAACCTTCACAACTTTGGGTTGTGCTTTGGCAATG	900
Db	882	TATTGCTATGTCAGCAGCAGGAGCAACCTTCACAACTTTGGGTTGTGCTTTGGCAATG	941
Qy	901	CCA 903	
Db	942	CCA 944	
RESULT 9			
US-08-937-399-23			
Sequence 23. Application US/08937399			
Patent No. 6286218			
GENERAL INFORMATION:			
APPLICANT: Levinson, Douglas A.			
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE			
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS			
NUMBER OF SEQUENCES: 37			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Pennie & Edmonds			
STREET: 1155 Avenue of the Americas			
CITY: New York			
STATE: New York			
COUNTRY: USA			
ZIP: 10036/2711			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Diskette			
COMPUTER: IBM Compatible			
OPERATING SYSTEM: DOS			
SOFTWARE: FastSeq Version 2.0			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/937,399			
FILING DATE:			
CLASSIFICATION:			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 08/609,583			
FILING DATE: 01-MAR-1996			
APPLICATION NUMBER: US 08/487,748			
FILING DATE: 07-JUN-1995			
APPLICATION NUMBER: US 08/398,633			
FILING DATE: 03-MAR-1995			
ATTORNEY/AGENT INFORMATION:			
NAME: Cortuzzi, Laura A.			
REGISTRATION NUMBER: 30,742			
REFERENCE/DOCKET NUMBER: 7853-048			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: 212-790-9090			
TELEFAX: 212-869-8864			
TELEX: 66141 PENNIE			
INFORMATION FOR SEQ ID NO: 23:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 2236 base pairs			
TYPE: nucleic acid			
STRANDEDNESS: single			
TOPOLOGY: linear			
MOLECULE TYPE: DNA			

```
; FEATURE:
; NAME/KEY : Coding Sequence
; LOCATION : 42...944
; OTHER INFORMATION: Human 200 gene nucleotide
; OTHER INFORMATION: sequence
; US-08-937-399-23

Query Match      100.0%; Score 903; DB 3; Length 226;
Best Local Similarity 100.0%; Pred. No. 2,36-297;
Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1 ATGTTTTCACATCTTCCTTGACTGTGTCTGTGCTGTGCTGTGCTGTACTACTTACAAG 60
Db      42 ATGTTTTACATCTTCCCTTGACTGTGTCTGTGCTGTGCTGTGCTGTCTACTTACAAG 101

QY      61 TCCTCAGAAGTGAATAATACAGACGGAGGTGGTCAGAAATGCCTATTCTGCCCTGTCTAC 120
Db      102 TCCTCAGAAGTGAATAATACAGACGGAGGTGGTCAGAAATGCCTATTCTGCCCTGTCTAC 161

QY      121 ACCCCAGCGCCCCAAGGGAACCTGTGCCCCGTCTGTGCGCAAGAGAGCTGTCTGTG 180
Db      162 ACCCCAGCGCCCCAAGGGAACCTGTGCCCCGTCTGTGCGCAAGAGAGCTGTCTGTG 221

QY      181 TTTGAATGTGGCAACGTGTGTCTCAGACTCATGAAAGGAATGTAAATTAATTGACATCC 240
Db      222 TTTGAATGTGGCAACGTGTGTCTCAGACTCATGAAAGGAATGTAAATTAATTGACATCC 281

QY      241 AGATTACTGGCTAAATGGGGATTTCCGCAAGAGAGATGTGTCCCTGACCATAGAGATGTG 300
Db      282 AGATTACTGGCTAAATGGGGATTTCCGCAAGAGAGATGTGTCCCTGACCATAGAGATGTG 341

QY      301 ACTCTAGCAGACAGTGGGATCTATCTGTCTGCCGATCCAAATCTCCAGGCTAATGAATGAT 360
Db      342 ACTCTAGCAGACAGTGGGATCTATCTGTCTGCCGATCCAAATCTCCAGGCTAATGAATGAT 401

QY      361 GAAAAATTTAACTGTAAAGTTGGTCAATCAAACGAGCAAGAGTCAACCCCTCACCGACTGTG 420
Db      402 GAAAAATTTAACTGTAAAGTTGGTCAATCAAACGAGCAAGAGTCAACCCCTCACCGACTGTG 461

QY      421 CAGAGAGACTTCACTGACGACCCTTTCCAAGATGCTTACCACAGGGGACATGACCACAGA 480
Db      462 CAGAGAGACTTCACTGACGACCCTTTCCAAGATGCTTACCACAGGGGACATGACCACAGA 521

QY      481 GAGACACAGACACTGGGGAGGCTCCCTGATATTAATCTTAACAACAATATCCAATTGGGCC 540
Db      522 GAGACACAGACACTGGGGAGGCTCCCTGATATTAATCTTAACAACAATATCCAATTGGGCC 581

QY      541 AATGAGTTACGGGACTCTTAGATTGGCCAAATGACTTACGGGACTCTGAGCAACCATCAGA 600
Db      582 AATGAGTTACGGGACTCTTAGATTGGCCAAATGACTTACGGGACTCTGAGCAACCATCAGA 641

QY      601 ATAAGCATCTACATGTGAGACAGAGGATCTGTGCTGGGCTGAGCTCTGTATTCTTCGGC 660
Db      642 ATAAGCATCTACATGTGAGACAGAGGATCTGTGCTGGGCTGAGCTCTGTATTCTTCGGC 701

QY      661 GCTTTAATTTTCAAAATGATATTTCTCATAGCAAAAGAAAGATACAGAAATTTAAGCCTCATC 720
Db      702 GCTTTAATTTTCAAAATGATATTTCTCATAGCAAAAGAAAGATACAGAAATTTAAGCCTCATC 761

QY      721 TCTTTGGCCAACTCTCCCTCCTCAGAGATTGGCAATGCAATGACAGAGGAATTCGCTCA 780
Db      762 TCTTTGGCCAACTCTCCCTCCTCAGAGATTGGCAATGCAATGACAGAGGAATTCGCTCA 821

QY      781 GAAGAAAACATCTATACATATGAAGAAGAACGTAATGSAAGTGAAGAGGCCAATGAGTAT 840
Db      822 GAAGAAAACATCTATACATATGAAGAAGAACGTAATGSAAGTGAAGAGGCCAATGAGTAT 881

QY      841 TATTGCTATGTCAAGCAGACAGGACGACCACTTCACACACTTTGGTGTGTGCTTTGCAATG 900
Db      882 TATTGCTATGTCAAGCAGACAGGACGACCACTTCACACACTTTGGTGTGTGCTTTGCAATG 941

QY      901 CCA 903
|||
```









QY	240	--CAGTACTGGCTAATGGGGATTTCCGAAAGGAGATGTGTCCCGACATATAGAAAT	237
Db	280	ACGAGATACAGCTTAAAGGCGATCTCAACAAAGAGATGTGTCTGATCATAAAGAAAT	339
QY	298	GTGACTCTAGACAGACAGTGGGATCTTACTGTGCGGATCCAAATCCGAGCATATGAAT	357
Db	340	GTGACTCTGATGACCATGGGACTTACTGTGTGAGGATTAAGTTCCCTGGTCTTATGAT	399
QY	358	GATGAAAAATTTAATCTGAAGTTGTGTATAAACAGCCAAAGTCAACCTCGACCGACT	417
Db	400	GATAAAAAATTTAGAACTGAATTTGACATCAAGCAGCCAAAGGTCACTCCAGTCAAGACT	459
QY	418	CTGACAGAGACCTTCACTGACGCTTCCAAAGATGGTTCCACCAAGGGGACATGGGCCA	477
Db	460	GCCCATGGGGACTCTTACTAGAGCTTCTCCAGAACCTTAAACACGAGAGAAATGG--T	516
QY	478	GCAGAGACACAGACACTGGGGAGCCTCCCTGATATATAATCTTAACAAATATCCACATGG	537
Db	517	TCAGAGACACAGACACTGGTGACCTCCATATATTAACATAGGAACAAAAATTTCCACATGG	576
QY	538	GCCATGAGTTACGGGACTCTTGAATTGGCCATGACTTACGGGACTCTGAGCAACCATC	597
Db	577	GCTGATGAATTTAA-----GGACTCTGAGAGAAACGATC	609
QY	598	AGATATGAGCATCTACATCGGAGAGGAGATCTGTCTGGGCTGGCTCGGCTTATCTTC	655
Db	610	AAGATCTGATCTACATCTTGAGATGGGAGATCTCTGCTGGGTTGACCTTGGACATTAAT	669
QY	658	GGCGCTTTAATTTTCAATGGTATTTCTCATAGCAAGAGAGATACAGATTTAAGCTTC	717
Db	670	GGTGTCTTATCTTAAATGGTATTTCTGTGAAGAAAAAGATTTATCGAATTTGAGCTT	729
QY	718	ATCTCTTTGGCCAACTCCCTCCCTAGAGATTTGGCAATTCAGTATGACAGAGGAATTCGC	777
Db	730	ATTACATGGCCAACTTGCTCTCCAGAGGGTTGGCAATTCAGAGACAGTCAAGATTCGC	789
QY	778	TCAGAGAAAAATCATCTATACCATTTGAAGAAACGTATATGATGAGAGAGCCCAATGAG	837
Db	790	TCTGAGAAAAATCTACACCATCGAGAGAACTTATGABATGAGAAATTCMAATGAG	849
QY	838	TATTATTTGCTATGTACAGACAGGAGCAACCTTCACAACTTTGGTGTGCTTTGCA	897
Db	850	TACTACTGCTAGTCAACAGCAGCGACCATCTGACCGGCTCTGACACTGCCACTTTAA	909
QY	898	ATGC 901	
Db	910	AGGC 913	

RESULT 14  
 US-08-480-070C-8  
 Sequence 8, Application US/08480070C  
 Patent No. 6066498  
 GENERAL INFORMATION:  
 APPLICANT: Levinson, Douglas A.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 NUMBER OF SEQUENCES: 24  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds LLP  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/480,070C

/	FILING DATE:	07-JUN-1995
/	CLASSIFICATION:	530
/	ATTORNEY/AGENT INFORMATION:	
/	NAME:	COTRUZZI, Laura A.
/	REGISTRATION NUMBER:	30,742
/	REFERENCE /DOCKET NUMBER:	7853-024
/	TELECOMMUNICATION INFORMATION:	
/	TELEPHONE:	(212) 790-9090
/	TELEFAX:	(212) 869-8864/9741
/	TELEX:	66141 PENNIE
/	INFORMATION FOR SEQ ID NO:	8:
/	SEQUENCE CHARACTERISTICS:	
/	LENGTH:	2710 base pairs
/	TYPE:	nucleic acid
/	STRANDEDNESS:	single
/	TOPOLOGY:	linear
/	MOLECULE TYPE:	DNA
/	FEATURE:	
/	NAME/KEY:	CDS
/	LOCATION:	40..885
/	US-08-480-070C-8	

  

Query Match	50.5%	Score 456;	DB 3;	Length 2710;
Best Local Similarity	72.0%;	Pred. No. 7.2e+145;	Mismatches 220;	Indels 33; Gaps 3
Matches 651; Conservative	0;			
OY	1	ATGTTTCACACTCTTCCCTTGAGCTGTGCTCGTGCCTGCTGCTACTACTTACAAAG	60	
Db	40	ATGTTTTAGGCTTTACCCTCAACCTGTCTCTGCTGCTGCTGCAACTTACTTCCAAAG	99	
OY	61	TCTCTCAGAAGTGGAAATPACAGACGGAGGTGGTCAAAAATGCCATCTGCCCTCTTAC	120	
Db	100	TCATTTGGAAGATGGTTTAATAAGTTGAGGTTGGTAAAAATGCCATCTGCCCTGCAGTTAC	159	
OY	121	ACCCGAGCGCGCCCAGGGAACCTGTGCGCGCTGTGGGGCAAAGAGCGCTGTCTGTG	180	
Db	160	ACTTACTCTAATCTGGGACACTTGTGCTTATGTCTGTGGGCAAGGATTTCTGTCTTGG	219	
OY	181	TTTGAATGTGGCAACGTGTGCTCAGACACTGATGAAAGGATGTGAATTAATTGACATC-	239	
Db	220	TCACAGTGTACCATGATGATTGCTCAGAACCTGATGAATAAATGTGACATATCAGAAATCC	279	
OY	240	--CAGATACCTGGCTTAATANGGGATTTCCGCAAAAGAGATGTGTCCCTGACATAGAGAT	297	
Db	280	AGCAGATACACGCTPAAGGGCGATCTCAACAAAGAGATGTGTCTGTATCATAAAGAT	339	
OY	298	GTCGACTTAGCAGACAGTGGGATCTACTGTGCCGATCCAATCCAGCATATGAAT	357	
Db	340	GTGACTGTGATGATCAATGGGACTACTGTCTGACGAGTACAGTTCCGTGTCTTATGAAT	399	
OY	358	GATGAAAAATTTAACTGGAAGTTGTCATCAACACGCCAAGTCAACCCCTGACCGACT	417	
Db	400	GATAAAAAATTGAACTGAAATTTGACATCAAGAGGCAAGGTCATCCACGCTCAGACT	459	
OY	418	CTGCAGAGAGCTTCACTGACGCTTTCCAAAGATGCTTACCAACAGGGGACATGGCCA	477	
Db	460	GCCCATGGGACTCTACTACAGCTTCTTCAGAAACCTTACCAAGAGAAATGG---T	516	
OY	478	GCAGAGACACAGACACTGGGGAGCCCTCCCTGATTAATAATCTPACAAATAATCACATVG	537	
Db	517	TCAGAGACACAGACACTGGTGACCTCCATTAATAATGAAACAAAATTTCCACATGG	576	
OY	538	GCCATAGTAGTACGGGACTCTAGATGGCCAATGACTTACGGGACTCTGGAGCAACATC	597	
Db	577	GCTGATGAAATTA-----GGACTCGAGAAACGATC	609	
OY	598	AGATAGGACATCTACATCGAGAGGAGATCTGTCTGGGCTGGCTGGCTTATCTTC	657	
Db	610	AGAGCTGTATCCACATTTGAGTGGAGTCTCTGCTGGTTGACCTTGACCTATATAT	669	
OY	658	GGCGCTTAATTTTCAAAATGTATTTCTCATRGCAGAAAGAAATPACGAATTTAAGCTCT	717	
Db	670	GGTGTCTTAATCCCTTAATGTGTATTTCTGTAAAGAAAAAGATTAATGAGTTTGAGCTCT	729	



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 12:17:01 ; Search time 1977.2 Seconds  
(without alignments)  
11099.996 Million cell updates/sec

Title: US-10-004-633-37

Perfect score: 903

Sequence: 1 atgtttcacatcttcctt.....gttgtcgttcgaatgcga 903

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estcov:\*  
6: em\_estp1:\*  
7: em\_estro:\*  
8: em\_hc1:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hc1:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	901	99.8	1201	13	BX417090	BX417090 BX417090
2	606	67.1	735	14	CA307836	CA307836 UI-H-FT1-
3	583	64.6	583	13	B0789118	B0789118 1040b02.x
4	568.2	62.9	863	14	CD520418	CD520418 AGENCOURT

5	525	58.1	554	13	B0581947	B0581947 1111a10.y
6	518.4	57.4	575	9	AU279844	AU279844
7	499.4	55.3	746	14	CD522575	CD522575 AGENCOURT
8	477.6	52.9	869	14	CD532271	CD532271 AGENCOURT
9	451.4	50.0	796	12	BI906710	BI906710 603064430
10	421	46.6	518	14	CB155731	CB155731 K-BST0214
11	397.4	44.0	572	10	AL709094	AL709094 DRFP2686F
12	396.6	43.9	837	10	BG545730	BG545730 602573003
13	349.8	38.7	519	12	BI834671	BI834671 603090448
14	345	38.2	1370	13	BQ278485	BQ278485 AGENCOURT
15	304.8	33.8	633	14	BY741594	BY741594 BY741594
16	277.8	30.8	677	14	BY748679	BY748679 BY748679
17	199.8	22.1	535	12	BI541253	BI541253 454905 MA
18	191	21.2	606	9	AV664631	AV664631 AV664631
19	150.6	16.7	429	9	AV664630	AV664630 AV664630
20	148	16.4	655	13	B0581620	B0581620 1111a10.x
21	147.6	16.3	557	12	BM255779	BM255779 517968 MA
22	137.2	15.2	350	13	BY185793	BY185793 BY185793
23	118.8	13.2	296	10	BB572107	BB572107 BB572107
24	111.6	12.4	634	13	B0238988	B0238988 603321725
25	111.6	12.4	694	9	AJ452225	AJ452225 AJ452225
26	111.6	12.4	776	9	AJ452245	AJ452245 AJ452245
27	111.6	12.4	807	9	AJ445463	AJ445463 AJ445463
28	102	11.3	754	13	B0320531	B0320531 603850370
29	100.4	11.1	900	13	B0346895	B0346895 603524895
30	100.2	11.1	544	13	BX264550	BX264550 BX264550
31	97	10.7	414	13	BQ484385	BQ484385 pmp1c.pko
32	91	10.1	820	13	B0207995	B0207995 603949365
33	82.4	9.1	615	13	BX265767	BX265767 BX265767
34	75	8.3	641	13	BQ400478	BQ400478 NISC mp09
35	75	8.2	828	14	CD254919	CD254919 AGENCOURT
36	74	8.2	568	12	BM261288	BM261288 da146d03
37	74	8.2	901	13	BU148837	BU148837 AGENCOURT
38	73.4	8.1	525	12	BI443222	BI443222 da138e11
39	72.6	8.0	722	9	AM604297	AM604297 IL3-CT021
40	72.4	8.0	754	14	CB938882	CB938882 IPGJX13
41	71.8	8.0	688	13	BU206552	BU206552 604159133
42	71	7.9	457	10	BE722324	BE722324 190515 MA
43	70.2	7.8	761	10	BG168159	BG168159 602341553
44	67.6	7.5	569	9	AJ454449	AJ454449 AJ454449
45	67.6	7.5	751	13	BX114328	BX114328 BX114328

## ALIGNMENTS

RESULT 1  
BX417090  
LOCUS  
DEFINITION BX417090 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DB0041YC05  
5-PRIME, mRNA sequence.  
ACCESSION BX417090  
VERSION BX417090.1 GI:30765665  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1201)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 2998.r for  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgl-bin/cluster.cgi?seq=CS0DB004AB03QPIcluster=2998.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Db 816 TCTTTGGCCAACTCTCCTCCCTCAGGATTGGCAATGCAGTAGCAGAGGGAAATTCGCTCA 875

Query Match	67.1%; Score 606; DB 14; Length 735;
-------------	--------------------------------------

	Best Local Similarity 89.2%; Pred. No. 3,1e-172;	Matches 700;	Conservative 0;	Mismatches 1;	Indels 84;	Gaps 1;
QY	119 ACACCCAGCCGCCCCCAGGGAACCTCGTCCGCTCTGTCTGGGGCAAGAGACCTGTCTCG	178				
Db	735 ACACCCAGCCGCCCCCAGGGAACCTCGTCCGCTCTGTCTGGGGCAAGAGACCTGTCTCG	676				
QY	179 TGTTTGAATGTGGCAACGTGTGCTCAGACCTGATGAAAGGATGTGAATTTATTTGGACAT	238				
Db	675 TGTNTGAATGTGGCAACGTGTGCTCAGACCTGATGAAAGGATGTGAATTTATTTGGACAT	616				
QY	239 CCAGATATCTGGCTTAATTGGGGATTTCGCCAAGAGAGATGTGCTCCCTGACCATAGAGATG	298				
Db	615 CCAGATATCTGGCTTAATTGGGGATTTCGCCAAGAGAGATGTGCTCCCTGACCATAGAGATG	556				
QY	299 TGACTTAGCAGACAGTGGGATCTTACTGTCTGCCGATCCAAATCCACGACATAATG	358				
Db	555 TGACTTAGCAGACAGTGGGATCTTACTGTCTGCCGATCCAAATCCACGACATAATG	496				
QY	359 ATGAAAAATTTAACTTGAAATTTGGTTCATCAAACCAAGGTCAACCTTGACCCGACTC	418				
Db	495 ATGAAAAATTTAACTTGAAATTTGGTTCATCAAACCAAGGTCAACCTTGACCCGACTC	464				
QY	419 TGCAGAGAGACTTCACTGCAGCCTTTCCAAAGATGCTTACCAACAGGGGACATGGCCAG	478				
Db	463 -----CCAG 460					
QY	479 CAGAGACACAGACACTGGGGAGCCTCCCTGTATTAATCTAACCAATATCCACATTTGG	538				
Db	459 CAGAGACACAGACACTGGGGAGCCTCCCTGTATTAATCTAACCAATATCCACATTTGG	400				
QY	539 CCAATGAGTTACGGGACTCTGATTTGGCCAAATGACTTACGGGACTCTGGAGCAACATCA	598				
Db	399 CCAATGAGTTACGGGACTCTGATTTGGCCAAATGACTTACGGGACTCTGGAGCAACATCA	340				
QY	599 GAATAGGATCTACATGAGAGACAGGAGATCTGTGCTGGGCTGCTCTCTATCTTCTGG	658				
Db	339 GAATAGGATCTACATGAGAGACAGGAGATCTGTGCTGGGCTGCTCTCTATCTTCTGG	280				
QY	659 GCGCTTTAAATTTCAAAATGATTTCTCATAGCAAAAGAGAAATACAGAAATTTAAGCTCA	718				
Db	279 GCGCTTTAAATTTCAAAATGATTTCTCATAGCAAAAGAGAAATACAGAAATTTAAGCTCA	220				
QY	719 TCTCTTTGGCCAACTCCCTCCCTCAGATTGGCAATGACATGACAGAGGGAATTCGCT	778				
Db	219 TCTCTTTGGCCAACTCCCTCCCTCAGATTGGCAATGACATGACAGAGGGAATTCGCT	160				
QY	779 CAGAAAGAAAATCTATTCATTCATTTGAAGGAACGTTATGAAGTGGAGAGCCCAATGATG	838				
Db	159 CAGAAAGAAAATCTATTCATTCATTTGAAGGAACGTTATGAAGTGGAGAGCCCAATGATG	100				
QY	839 ATTATTCGTATGTGACGACAGAGGAGCAACCTCCAACTTTGGGTTGTGCGCTTTGGCA	898				
Db	99 ATTATTCGTATGTGACGACAGAGGAGCAACCTCCAACTTTGGGTTGTGCGCTTTGGCA	40				
QY	899 TGCCA 903					
Db	39 TGCCA 35					
RESULT 3						
LOCUS	BUT89118					
DEFINITION	ic40b02.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:6128595					
ACCESSION	BUT89118					
VERSION	BUT89118.1					
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catartini; Homnidae; Homo.					

REFERENCE	1 (bases 1 to 583)
AUTHORS	Melton,D., Brown,T., Keny,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Merra,M., Pape,D., Wylie,T., Martin,J., Blistan,A., Schiller,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas ,M., Gibbons,M., McCann,R., Cole,R., Taagereshvilli,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE	Endocrine Pancreas Consortium
JOURNAL	Unpublished
COMMENT	Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue Endocrine Pancreas Consortium Harvard University, Howard Hughes Medical Institute Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138 Tel: 617-495-1812 Fax: 617-495-8557 Email: dmelton@diobp.harvard.edu
FEATURES	Library was constructed by Dr. J. Ferrer in vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu) Possible reversed clone: similarity on wrong strand Seq primer: -40UP from Gibco High quality sequence stop: 460. Location/Qualifiers
SOURCE	1. 583 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6128595" /tissue_type="Insulinoma" /lab_host="DH10B (phage-resistant)" /note="Organ: pancreas; Vector: pBluescript SK-; Site 1: XhoI; Site 2: EcoRI; Constructed with lambda ZapII system (Stratagene) by Dr. J. Ferrer. in vivo mass-excised to pBluescript SK- by Dr. H. Inoue Following the Washington University protocol (http://genome.wustl.edu/est/lambda_protocol.shtml). Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: This is a Washington University Pancreas EST project library."
BASE COUNT	152 a 152 c 148 g 131 t
ORIGIN	
Query Match	64.6%; Score 583; DB 13; Length 583;
Best Local Similarity	100.0%; Pred. No. 2,6e-165;
Matches 583; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Dy	7 TCACATCTCCCTTGACTGTCGTCTCGTGCTGCTGACTACTTAAGAAGTCCCA 65
Db	1 TCACATCTCCCTTGACTGTCGTCTCGTGCTGCTGACTACTTAAGAAGTCCCA 60
Dy	67 GAAGTGAATAACAAGCGAGGTGCTCAGATTGCCATATGCCTTCCTTCAACCCCA 126
Db	61 GAATGGAATACAGAGCGAGGTGCTCAGATTGCCATATGCCTTCCTTCAACCCCA 120
Dy	127 GCCGCCCCOAGGAACTCTGTCGCCCTTGTCTGGGGCAAAGAGCTCTGTGTTTGAA 186
Db	121 GCCGCCCCOAGGAACTCTGTCGCCCTTGTCTGGGGCAAAGAGCTCTGTGTTTGAA 180
Dy	187 TGTGGCAACGTGTCCTCAGGACTGATGAAGAAAGGATGTGAATTATTGGACATCCAGATAC 246
Db	181 TGTGGCAACGTGTCCTCAGGACTGATGAAGAAAGGATGTGAATTATTGGACATCCAGATAC 240
Dy	247 TGCGTAATATGGGAGTTTCGCAAGAGAGATGTGTCCCTGACCATAGAGAAATGTGACTCTA 306
Db	241 TGCGTAATATGGGAGTTTCGCAAGAGAGATGTGTCCCTGACCATAGAGAAATGTGACTCTA 300
Dy	307 GCAGACATGGGATCTACTGTCTGCCGATCCAATCCAGGCATATGAATGATGAAAAA 366

Db 301 GGAGACAGTGGAGTCTACTGCTGCGGATCCAGCAATCCAGCAATATGATGATGAAAA 360  
QY 367 TTTTAACTGAAGTTGGTCATCAAAACGACCAAGTCACCCCTGACCCGACTCTGACAGAGA 426  
Db 361 TTTTAACTGAAGTTGGTCATCAAAACGACCAAGTCACCCCTGACCCGACTCTGACAGAGA 420  
QY 427 GACTTACTGACAGCTTTTCCAGAGATGCTTACGACAGAGGGAGCATGGCCGACAGAGAGA 486  
Db 421 GACTTACTGACAGCTTTTCCAGAGATGCTTACGACAGAGGGAGCATGGCCGACAGAGAGA 480  
QY 487 CAGACACTGGGAGGAGCTCCCTGATATATATTAATCTAACAACAATATCAATTGGCCAAATGAG 546  
Db 481 CAGACACTGGGAGGAGCTCCCTGATATATATTAATCTAACAACAATATCAATTGGCCAAATGAG 540  
QY 547 TTTAGGGAGCTTGTGATTTGGCCATGACTTACGAGGACTCTGGAG 589  
Db 541 TTTAGGGAGCTTGTGATTTGGCCATGACTTACGAGGACTCTGGAG 583  
RESULT 4  
CD520418 863 bp mRNA linear EST 06-JUN-2003  
LOCUS AGENCOURT 14354411 NIH MGC 191 Homo sapiens cDNA clone  
DEFINITION IMAGE:30411529 5', mRNA sequence.  
ACCESSION CD520418  
VERSION CD520418.1 GI:31452136  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 863)  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished  
JOURNAL Office of Cancer Genomics  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: ggaabs-remail.nih.gov  
Tissue Procurement: Narayan Bhat  
cDNA Library Preparation: Clontech Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: NDCM203 row: h column: 02  
High quality sequence stop: 509.  
Location/Qualifiers  
1..863  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30411529"  
/issue\_type="Pooled"  
/lab\_host="DH10B (TI phage-resistant)"  
/clone\_lib="NIH\_MGC\_191"  
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);  
Site 2: SfiI (ggccgcctggcc); Library is oligo-dr primed  
and directionally cloned. PBMC - Peripheral Blood  
Mononuclear Cells. RNA was pooled from 3/6hour stimulation  
with FMA adn Ionomycin. 5' and 3' adaptors were used in  
cloning as follows: 5' adaptor sequence:  
5'-CACGGCCATTATGCGC-3' and 3' adaptor sequence:  
5'-ATTCTAGGAGCCGAGGCGGCGCATG-dr(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.69  
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA). Note: this is a NIH MGC Library."

BASE COUNT  
ORIGIN

221 a 222 c 203 g 215 t 2 others

Query Match 62.9%; Score 568.2; DB 14; Length 863;  
Best Local Similarity 96.5%; Fred. No. 1e-160;  
Matches 612; Conservative 0; Mismatches 19; Indels 3; Gaps 3;  
QY 1 ATGTTTACATCTTCCCTTTGACATGTCGTCGCTGCTGCTACTACTTACAG 60  
Db 59 ATGTTTACATCTTCCCTTTGACATGTCGTCGCTGCTGCTACTACTTACAG 118  
QY 61 TCCTCAGAGTGAATACAGAGCGAGAGTGGTCAAGATGCTATCTGCTCTTCTAC 120  
Db 119 TCCTCAGAGTGAATACAGAGCGAGAGTGGTCAAGATGCTATCTGCTCTTCTAC 178  
QY 121 ACCCCAGCCGCCCCAGAGAACTCTGTCGCCCTGTCGTCGGGCAAGAGACCTCTCTG 180  
Db 179 ACCCCAGCCGCCCCAGAGAACTCTGTCGCCCTGTCGTCGGGCAAGAGACCTCTCTG 238  
QY 181 TTTGAATGTGGCAACGTGTGCTCAGAGACTGATGAAGGATGTGAATTATTGACATCC 240  
Db 239 TTTGAATGTGGCAACGTGTGCTCAGAGACTGATGAAGGATGTGAATTATTGACATCC 298  
QY 241 AGATCTGCTTAATATGGGATTTTCGCAAGAGATGTGTCTCTGACATAGAGATGTG 300  
Db 299 AGATCTGCTTAATATGGGATTTTCGCAAGAGATGTGTCTCTGACATAGAGATGTG 358  
QY 301 ACTCTGACAGACAGTGGGATCTACTGTCGCCGATTCGAATCCAGGCAATATGAT 360  
Db 359 ACTCTGACAGACAGTGGGATCTACTGTCGCCGATTCGAATCCAGGCAATATGAT 418  
QY 361 GAAAAATTTAACTGAAGTTGGTCATCAAAACGACCAAGGTCACCCCTGACCGACTCTG 420  
Db 419 GAAAAATTTAACTGAAGTTGGTCATCAAAACGACCAAGGTCACCCCTGACCGACTCTG 478  
QY 421 CAGAGAGACTTCACTGACGCTTTTCCAGAGATGCTTACCAAGGAGATGGCCAGCA 480  
Db 479 CAGAGAGACTTCACTGACGCTTTTCCAGAGATGCTTACCAAGGAGATGGCCAGCA 538  
QY 481 GAGACACAGACACT -GGGAGGCTCCCTGATTAATCTTAACAACAATATTCGATGGC 539  
Db 539 GAGACACAGACACTGGGAGGCTCCCTGATTAATCTTAACAACAATATTCGATGGC 598  
QY 540 CAATGAGTTAC -GGGACTCTAGATTGGCCAAATGACTTACGGGACTCTG -AGCAACATC 597  
Db 599 CAATGAGTTACGGGAGCTCTAGATTGGCCAAATGACTTACGGGACTCTGAGCAACATC 658  
QY 598 AGAATAGGATCTTACATCGAGAGGATCTGTG 631  
Db 659 AAAATAGGATCTTACCTTCGACAGAGGATCTGTG 692  
RESULT 5  
BQ581947 554 bp mRNA linear EST 20-JUN-2002  
LOCUS BQ581947  
DEFINITION 1111a10.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6029490  
5' similar to TR:054947 054947 KIDNEY INJURY MOLECULE-1 PRECURSOR  
1..554  
/RNA sequence.  
ACCESSION BQ581947  
VERSION BQ581947.1 GI:21494843  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 554)  
REFERENCE Melton, D., Brown, J., Kenly, G., Permut, A., Lee, C., Kaestner, K.,  
Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blaisdin, A.,  
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas  
M., Gibbons, W., McCann, R., Cole, R., Tsagarisvilli, R., Williams, F.,  
Jackson, Y., and Bowers, Y.  
Endocrine Pancreas Consortium  
TITLE Unpublished  
JOURNAL Other\_ESTS: 1111a10.x1  
COMMENT





DB 356 ACTTGGACAGACGTGGATCTACTGCTCCGGATCCAAATCCAGGCATTAATGAT 415  
QY 361 GAAAAATTTAACTGAAAGTTGTCATCAACACGACAGGTCAACCCCTGACCGACTCTG 420  
DB 416 GAAAAATTTAACTGAAAGTTGTCATCAACACGACAGGTCAACCCCTGACCGACTCTG 475  
QY 421 CAGAGAGACTTCACTGACGCTTTCCAAAGATGCTTACACGACGAGGACATGGCCACGA 480  
DB 476 CAGAGAGACTTCACTGACGCTTTCCAAAGATGCTTACACGACGAGGACATGGCCACGA 535  
QY 481 GAGACACAGACACTGGGAGGAGCTCCCGATATTAATCTAA 520  
DB 536 GAGACACAGACACTGGGAGGAGCTCCCGATATTAATCTAA 575

RESULT 7  
CD522575 746 bp mRNA linear EST 06-JUN-2003  
LOCUS AGENCOURT\_1435479 NIH MGC\_191 Homo sapiens cDNA clone  
DEFINITION IMAGE:30411065 5', mRNA sequence.  
ACCESSION CD522575  
VERSION CD522575.1 GI:31454293  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 746)  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Narayan Bhat  
cDNA Library Preparation: Clontech Laboratories, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: NDCM202 row: d column: 18  
High quality sequence stop: 592.  
Location/Qualifiers  
1..746  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30411065"  
/issue\_type="Pooled"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH MGC 191"  
/note="Vector: pDR-Lib, Site 1: SfiI (ggccatctggcc);  
Site 2: SfiI (ggccgcctcgcc); Library is oligo-dT primed  
and directionally cloned. PBMC - Peripheral Blood  
Mononuclear Cells. RNA was pooled from 3/6hour stimulation  
with PMA adn Ionomycin. 5' and 3' adaptors were used in  
cloning as follows: 5' adaptor sequence:  
5'-CACGCGCATTTGSCC-3' and 3' adaptor sequence:  
5'-ATTCTAGAGCCGAGGCGCGACATG-dT(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.69  
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA). Note: this is a NIH MGC Library."

BASE COUNT 191 a 185 c 192 g 177 t 1 others  
ORIGIN

Query Match 55.3%; Score 499.4; DB 14; Length 746;  
Best Local Similarity 97.6%; Pred. No. 7e-140;

Matches 518; Conservative 0; Mismatches 11; Indels 2; Gaps 1;  
QY 1 ATGTTTCAACATCTTCCCTTTGACTGTGTCTGCTGCTGCTGCTACTTACAAAG 60  
DB 143 ATGTTTCAACATCTTCCCTTTGACTGTGTCTGCTGCTGCTGCTACTTACAAAG 202  
QY 61 TCCTCAGAGTGGAAATACAGAGCGGAGGTGGTCCAGAAAGCTATGCTGCTCTCAG 120  
DB 203 TCCTCAGAGTGGAAATACAGAGCGGAGGTGGTCCAGAAAGCTATGCTGCTCTCAG 262  
QY 121 ACCCCAGCCGCCCCAGAGGAACCTGCTGCCGCTGCTGCTGGGCAAGAGACCTGTCTG 180  
DB 263 ACCCCAGCCGCCCCAGAGGAACCTGCTGCCGCTGCTGCTGGGCAAGAGACCTGTCTG 322  
QY 181 TTTGAATGTGGCAACGTGTGCTCAGAGACTGATGAAGGAGTGTGAATTATGGACATCC 240  
DB 323 TTTGAATGTGGCAACGTGTGCTCAGAGACTGATGAAGGAGTGTGAATTATGGACATCC 382  
QY 241 AGATTAAGTGAATGAGGAGATTTCCGCAAGAGAGATGTGCTCCCTGACATAGAGATGTG 300  
DB 383 AGATTAAGTGAATGAGGAGATTTCCGCAAGAGAGATGTGCTCCCTGACATAGAGATGTG 442  
QY 301 ACTCTAGCAGACAGTGGGATCTACTGCTGCCGATCCAAATCCAGGACATTAATGAATGAT 360  
DB 443 ACTCTAGCAGACAGTGGGATCTACTGCTGCCGATCCAAATCCAGGACATTAATGAATGAT 502  
QY 361 GAAAAATTTAACTGAAAGTTGTCATCAACACGACAGGTCAACCCCTGACCGACTCTG 420  
DB 503 GAAAAATTTAACTGAAAGTTGTCATCAACACGACAGGTCAACCCCTGACCGACTCTG 562  
QY 421 CAGAGAGACTTCACTGACGCTTTCCAAAGATGCTTACACGACGAGGACATGGCCACGA 480  
DB 563 CAGAGAGACTTCACTGACGCTTTCCAAAGATGCTTACACGACGAGGACATGGCCACGA 622  
DB 481 GAGACACAGACACTGGGAGGAGCTCCCTTGATTAATCTAAACAAATAT 529  
DB 623 GAGACACAGACACTGGGAGGAGCTCCCTTGATTAATCTAAACAAATAT 673

RESULT 8  
CD523271 869 bp mRNA linear EST 06-JUN-2003  
LOCUS AGENCOURT\_1435479 NIH MGC\_191 Homo sapiens cDNA clone  
DEFINITION IMAGE:30410457 5', mRNA sequence.  
ACCESSION CD523271  
VERSION CD523271.1 GI:31454989  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 869)  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Narayan Bhat  
cDNA Library Preparation: Clontech Laboratories, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: NDCM200 row: k column: 10  
High quality sequence stop: 435.  
Location/Qualifiers  
1..869  
/organism="Homo sapiens"  
/mol\_type="mRNA"

FEATURES  
source



QY 528 ATCCA 532  
Db 791 TTCCA 795

RESULT 10  
LOCUS CB155731 518 bp mRNA linear EST 29-JAN-2003  
DEFINITION K-EST0214151 B2N807043 Homo sapiens cDNA clone B2N807043-7-D07 5', mRNA sequence.  
ACCESSION CB155731  
VERSION CB155731  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 518)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,D.M., Park,H.S., Kim,S. and Kim,Y.S.  
21C Frontier Korean EST Project 2001  
TITLE Unpublished  
JOURNAL Contact: Kim YS  
COMMENT Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongjung@mail.kribb.re.kr  
Plate: 7 row: D column: 07  
High quality sequence stop: 518.  
Location/Qualifiers  
1..518  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="B2N807043-7-D07"  
/sex="M"  
/lab\_host="Top10F"  
/clone\_1id="B2N807043"  
/note="Organ: Brain; Vector: pCNS-D2; Site 1: EcoRI; Site 2: NotI. The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 128 a 128 c 132 g 130 t

ORIGIN

Query Match 46.6%; Score 421; DB 14; Length 518;  
Best Local Similarity 100.0%; Pred. No. 3,4e-116;  
Matches 421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTTCACATCTTCCTTGAAGTGTCTGCTGCTGCTACTACTTACAGG 60  
|||||  
Db 98 ATGTTTTCACATCTTCCTTGAAGTGTCTGCTGCTGCTACTACTTACAGG 157  
|||||

QY 61 TCCTCAGAGTGAATACAGAGCGAGTGGTCAGATGCTATCTGCTGCTTCTAC 120  
|||||  
Db 158 TCCTCAGAGTGAATACAGAGCGAGTGGTCAGATGCTATCTGCTGCTTCTAC 217  
|||||

QY 121 ACCCCAGCCGCCAGGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
|||||

Db 218 ACCCAGCCGCCAGGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 277  
|||||

QY 181 TTGATGTGGCAACGTGTGCTCAGACTGATGAAGGATGTGATTTATGACATCC 240  
|||||

Db 278 TTGATGTGGCAACGTGTGCTCAGACTGATGAAGGATGTGATTTATGACATCC 337  
|||||

QY 241 AGATCTGCTTAATGAGGATTTCCGCAAGAGATGTCTCCCTGACCATAGAGATG 300  
|||||

Db 338 AGATCTGCTTAATGAGGATTTCCGCAAGAGATGTCTCCCTGACCATAGAGATG 397  
|||||

QY 301 ACTCTGACAGACAGTGGATCTACGTGCGGATCCAGATCCAGCATTAATGATGAT 360  
|||||

Db 398 ACTCTGACAGACAGTGGATCTACGTGCGGATCCAGATCCAGCATTAATGATGAT 457  
|||||

QY 361 GAAAAATTAACTGAAGTTGTGATCAACAGCCAGATGACCCCTGACCGACTGTG 420  
|||||

Db 458 GAAAAATTAACTGAAGTTGTGATCAACAGCCAGATGACCCCTGACCGACTGTG 517  
|||||

QY 421 C 421  
|

Db 518 C 518

RESULT 11  
LOCUS AL709094 572 bp mRNA linear EST 22-MAR-2002  
DEFINITION DKFZp686F1557\_r1 686 (synonym: h1cc3) Homo sapiens cDNA clone  
ACCESSION DKFZp686F1557 5', mRNA sequence.  
VERSION AL709094  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 572)  
AUTHORS Ottenwaelder,B., Obermaier,B., Mewes,W., Mewes,H.W., Weil,B. and Wiemann,S.  
EST (Ottenwaelder,B., Obermaier,B., Mewes,H.W., Weil,B. and Wiemann,S.)  
TITLE Unpublished  
JOURNAL Contact: Ottenwaelder B  
COMMENT MIPS  
Ingolstaedter Landstr.1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by MedGenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. No sl sequence available.  
This clone (DKFZp686F1557) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1..572  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686F1557"  
/issue\_type="human skeletal muscle"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1id="686 (synonym: h1cc3)"  
/note="Vector: pTriplEx2; Site\_1: SfiI; Site\_2: SfiIb; cDNA-collection"

BASE COUNT 141 a 137 c 152 g 142 t

ORIGIN

Query Match 44.0%; Score 397.4; DB 9; Length 572;  
Best Local Similarity 99.7%; Pred. No. 5,2e-109;  
Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT	12
BG545730	
LOCUS	837 bp mRNA linear EST 04-APR-2001
DEFINITION	602573003r1 NIH_MGC_77 Homo sapiens CDNA clone IMAGE:4701288 5', mRNA sequence.
ACCESSION	BG545730
VERSION	BG545730.1 GI:13544395
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 837)	NIH-MGC	<a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
	National Institutes of Health,	Mammalian Gene Collection (MGC)		
	Unpublished			
Contact:	Robert Strausberg, Ph.D.			
Email:	<a href="mailto:cgaabbs-remail.nih.gov">cgaabbs-remail.nih.gov</a>			
Tissue Procurement:	ClONTECH laboratories, Inc.			
cDNA Library Preparation:	ClONTECH laboratories, Inc.			
cDNA Sequencing by:	The I.M.A.G.E. Consortium (LNLN)			
DNA Distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:			
<a href="http://image.lnl.gov">http://image.lnl.gov</a>				
Plate:	LtCM1536 row: e column: 01			
High quality sequence stop:	722.			
Location/Qualifiers	I . 837			

```

FEATURES
    source
        location/Qualifiers
            1..837
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4701288"
                /lab_host="DH10B (T1 phage-resistant)"
                /clone_id="NH MGC 77"
                /note="Organ: lung; Vector: pDMR-LIB (Clontech); Site: 1:
                    SfiI (ggcgccctcgcc); Site_: SfiI (ggccatcagcc); 5' and
                    3' adaptors were used in cloning as follows: 5' adaptor
                    sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor sequence
                    5'-ATTCTAGAGCCGAGCGCCGACATG-dT (30)BN-3' (where B = A,
                    C, or G and N = A, C, G, or T). Average insert size 1.9
                    kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                    by PCR. This library was enriched for full-length clones
                    and was constructed by Clontech Laboratories (Palo Alto,

```

Query Match	43.9%	Score 396.6	DB 10	length 837
Best Local Similarity	97.8%	Pred. No. 1,2e-108		
Matches 402; Conservative	0	Mismatches 9	Indels 0	Gaps 0

OY	1	ATGTTTCAACATCTTCCCTTGAATGTCCTCGTGCCTGCTGCTGCTATCTATCAAGG	60
Dd	111	ATGTTTCAACATCTTCCCTTGAATGTCCTCGTGCCTGCTGCTGCTATCTATCAAGG	170
OY	61	TCCCTCAGAAATGGAATA CAGAGCCGAGTGGTCAGATGCTATCTGCCCCGTCTTAC	120
Dd	171	TCCCTCAGAAATGGAATA CAGAGCCGAGTGGTCAGATGCTATCTGCCCCGTCTTAC	230
OY	121	ACCCGAGCCGCTCCCAAGGGAACCTTGTCGCTGCTGCTGCGGCAAGAGCCTGCTGTG	180
Dd	231	ACCCGAGCCGCTCCCAAGGGAACCTTGTCGCTGCTGCGGCAAGAGCCTGCTGTG	290
OY	181	TTTGAATGTGGCAACGTGTGCTCAGACGTGATGAAGGGAATGGAATATTGGACATCC	240
Dd	291	TTTGAATGTGGCAACGTGTGCTCAGACGTGATGAAGGGAATGGAATATTGGACATCC	350
OY	241	AGATACCTGCTAAATGGGGATTTTCCGCAAGAGAAATGTGCTCTGACCATAGAGATGTG	300
Dd	351	AGATACCTGCTAAATGGGGATTTTCCGCAAGAGAAATGTGCTCTGACCATAGAGATGTG	410
OY	301	ACTCTAGCAGACAGTGGGATTTACTGCTGCCGGAATCCAAATCCAGGCATATGATGAT	360
Dd	411	ACTCTAGCAGACAGTGGGATTTACTGCTGCCGGAATCCAAATCCAGGCATATGATGAT	470
OY	361	GAAGAAATTTAACTGAAGTGGTCATCAAAACAGACAAAGTCACCCCCCA	411
Dd	471	GAAGAAATTTAACTGAAGTGGTCATCAAAACAGAGTAGTGGACATTTTGGCA	521

RESULT	13
B1834671	
LOCUS	B1834671
DEFINITION	B1834671 519 bp mRNA linear EST 04-OCT-2001
ACCESSION	G03090446F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5229539 5'
VERSION	B1834671
KEYWORDS	B1834671.1 GI:15946208
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL	1 (bases 1 to 519)
COMMENT	NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D.

```

FEATURES
source
    Tissue Procurement: Life Technologies, Inc.
    cDNA Library Preparation: Life Technologies, Inc.
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MSC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LNLN at:
    http://image.lnl.gov
    Plate: L1AM11577 row: 0 column: 12
    High quality sequence stop: 496
    Location/Qualifiers
        1..519
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5229539"
            /lab_host="DH10B"
            /clone_id="NH_MGC_120"
            /note="Organ: pooled pancreas and spleen; Vector:

```





Mon Nov 24 13:30:28 2003

Job time : 1981.2 secs

us-10-004-633-37.rst

Page 12



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 11:56:41 ; Search time 10120.7 Seconds  
(without alignments)  
10954.308 Million cell updates/sec

Title: US-10-004-633-8  
Perfect score: 2710  
Sequence: 1 ngcgcaccaccgcgcgcga.....aaaaaaaaagggcgccgc 2710

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb Da: \*  
2: gb Hcg: \*  
3: gb In: \*  
4: gb Om: \*  
5: gb Ov: \*  
6: gb Pat: \*  
7: gb Ph: \*  
8: gb Pl: \*  
9: gb Pr: \*  
10: gb Ro: \*  
11: gb Sgs: \*  
12: gb Sy: \*  
13: gb Un: \*  
14: gb Vi: \*  
15: em Ba: \*  
16: em Fun: \*  
17: em Hum: \*  
18: em In: \*  
19: em Mu: \*  
20: em Om: \*  
21: em Or: \*  
22: em Ov: \*  
23: em Pat: \*  
24: em Ph: \*  
25: em Pl: \*  
26: em Ro: \*  
27: em Sgs: \*  
28: em Un: \*  
29: em Vi: \*  
30: em Hcg\_hum: \*  
31: em Hcg\_inv: \*  
32: em Hcg\_other: \*  
33: em Hcg\_mus: \*  
34: em Hcg\_pln: \*  
35: em Hcg\_rod: \*  
36: em Hcg\_mam: \*  
37: em Hcg\_vrt: \*  
38: em Sy: \*  
39: em Htgo\_hum: \*  
40: em Htgo\_mus: \*  
41: em Htgo\_other: \*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2706.6	99.9	2710	6	AR143561
2	2706.6	99.9	2710	6	AR168930
3	2706.6	99.9	2710	6	AR232678
4	2706.6	99.9	2710	6	AR232678
5	2624.6	96.8	2725	10	AF450241
6	1886.6	69.6	223030	10	AL669948
7	1144.8	42.2	177087	2	AC119505
8	1144.8	42.2	233370	2	AC098539
9	844.4	31.2	862	10	AF39831
10	843	31.1	843	6	AR143567
11	843	31.1	843	6	AR168936
12	843	31.1	843	6	AR232684
13	499.8	18.4	1772	6	BD157433
14	499.8	18.4	1772	9	AK027334
15	499.8	18.4	2236	6	AR143568
16	499.8	18.4	2236	6	AR168937
17	499.8	18.4	2236	6	AR232685
18	498.2	18.4	2320	9	AF251707
19	485.8	17.9	1116	9	AF450242
20	485.8	17.9	1116	9	AF450243
21	456	16.8	903	6	AR143581
22	456	16.8	903	6	AR168950
23	456	16.8	903	6	AR232698
24	456	16.8	906	6	AY069944
25	336.4	12.4	769	6	BD146047
26	263.4	9.7	529	6	BD060055
27	255.6	9.4	177087	2	AC119505
28	255.6	9.4	233370	2	AC098539
29	231.6	8.5	1012	9	BC020843
30	219.8	8.1	438	6	BD076086
31	208.8	7.7	223030	10	AL669948
32	196.2	7.2	228827	2	AC135694
33	189.4	7.0	54375	9	AC011377
34	189.4	7.0	54375	9	AC011377
35	104.2	4.2	188225	2	AC108587
36	103.6	3.8	213613	2	AC110244
37	103.6	3.8	269920	2	AC094262
38	102.6	3.8	175542	2	AC112760
39	102.6	3.8	211694	2	AC141141
40	100	3.8	216060	2	AC097235
41	96.8	3.7	190756	2	AC107854
42	96.4	3.6	257333	2	AC137722
43	95.2	3.6	236117	2	AC106381
44	94.4	3.5	203206	2	AC109866
45	92.4	3.5	188016	2	AC121256
		3.4	76318	2	AC100517

#### ALIGNMENTS

RESULT 1  
AR143561  
LOCUS  
DEFINITION Sequence 8 from patent US 6204371.  
ACCESSION AR143561  
VERSION AR143561.1 GI:15104847  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2710)  
AUTHORS Levinson,D.Adam.  
TITLE Compositions and methods for the treatment and diagnosis of immune disorders  
JOURNAL Patent: US 6204371-A 8 20-MAR-2001;

Pred. No. is the number of results predicted by chance to have a

FEATURES	Location/Qualifiers
source	1. .2710
BASE COUNT	789 a 592 c 644 g 678 t 7 others
ORIGIN	
Query Match	99.9%; Score 2706.6; DB 6; Length 2710;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2709; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Db	2 GTCGACCCACGGGTCCCGGATTTCCCTCCCAAGTACTATGTTTTCAGGTCCTTACCCTCA 61
Db	2 GTCGACCCACGGGTCCCGGATTTCCCTCCCAAGTACTATGTTTTCAGGTCCTTACCCTCA 61
Qy	62 ACTGTGTCTGTCTGTCTGTCAACTACTACTTGCAGAGTCACTTGAAGATGGTTATTAAG 121
Db	62 ACTGTGTCTGTCTGTCTGTCAACTACTACTTGCAGAGTCACTTGAAGATGGTTATTAAG 121
Qy	122 TTGAGGTTGGTAAAAATGCTTATCTGCGCTGCAGATTACACTTACTTACCTTGGGACAC 181
Db	122 TTGAGGTTGGTAAAAATGCTTATCTGCGCTGCAGATTACACTTACTTACCTTGGGACAC 181
Qy	182 TTGTGCCATATGTCGTGGGGCAGAGGATTCGTGCTTGGCAAGTGTCCCATGAGTGC 241
Db	182 TTGTGCCATATGTCGTGGGGCAGAGGATTCGTGCTTGGCAAGTGTCCCATGAGTGC 241
Qy	242 TCAGAACTGATGAAAAATGTGACATATCAGAAAAATCCAGCATACGATCAAGCTTAAAGGCG 301
Db	242 TCAGAACTGATGAAAAATGTGACATATCAGAAAAATCCAGCATACGATCAAGCTTAAAGGCG 301
Qy	302 ATCTCAACAAAGAGATGTGTCTTGATCATTAAGAATGTGACTCTGATGACCATGGGA 361
Db	302 ATCTCAACAAAGAGATGTGTCTTGATCATTAAGAATGTGACTCTGATGACCATGGGA 361
Qy	362 CCTACTCGTCGAGGATATACAGTTCCTCGGTCTTAATGAATGAATTAATAAATTTAGACCTGAAT 421
Db	362 CCTACTCGTCGAGGATATACAGTTCCTCGGTCTTAATGAATGAATTAATAAATTTAGACCTGAAT 421
Qy	422 TAGACATCAAAACAGCCAGCTCACTCCAGCTCAGACTGCGATGGGAGCTTACTACTAG 481
Db	422 TAGACATCAAAACAGCCAGCTCACTCCAGCTCAGACTGCGATGGGAGCTTACTACTAG 481
Qy	482 GTTCTCCAAAGACCTTAACCAACGAGAGAAATGTTTCAAGACACACACACTGCTGACCC 541
Db	482 GTTCTCCAAAGACCTTAACCAACGAGAGAAATGTTTCAAGACACACACTGCTGACCC 541
Qy	542 TCCATATATACAAATGGAACAAAAATTTCCACATGGGCTGATTAATAATTAAGGACTCTGGAG 601
Db	542 TCCATATATACAAATGGAACAAAAATTTCCACATGGGCTGATTAATAATTAAGGACTCTGGAG 601
Qy	602 AAACGATCGAAGCTGTATCCACATTTGAGTGGGAGTCTCTGCTGGGTGACCTTGGCAC 661
Db	602 AAACGATCGAAGCTGTATCCACATTTGAGTGGGAGTCTCTGCTGGGTGACCTTGGCAC 661
Qy	662 TTATCATTTGGTCTTAAATCTTAAATGGTATTCCTGTAAGAAAAAGATTTATCGAGTT 721
Db	662 TTATCATTTGGTCTTAAATCTTAAATGGTATTCCTGTAAGAAAAAGATTTATCGAGTT 721
Qy	722 TTAGCGCTTATTAACCTGGGCAACTTGCCTCCAGAGAGGTGTGGCAATGAGAGAGACGCA 781
Db	722 TTAGCGCTTATTAACCTGGGCAACTTGCCTCCAGAGAGGTGTGGCAATGAGAGAGACGCA 781
Qy	782 GGAATTCGCTCTGAGGAAAAATATCTACACCATGAGAGAAAGTATATGAATGGAGAAATT 841
Db	782 GGAATTCGCTCTGAGGAAAAATATCTACACCATGAGAGAAAGTATATGAATGGAGAAATT 841
Qy	842 CAAATGAGTACTATGCTCACTGTCACACAGCAGCAGGACATCTGACCGCTCTGAGCTGCC 901
Db	842 CAAATGAGTACTATGCTCACTGTCACACAGCAGCAGGACATCTGACCGCTCTGAGCTGCC 901
Qy	902 ACTTTTAAAGGCTCGGCTTCAATTTCTGACTTTTGGATTTTCCCTTTTGTGAAAACTATGTG 961
Db	902 ACTTTTAAAGGCTCGGCTTCAATTTCTGACTTTTGGATTTTCCCTTTTGTGAAAACTATGTG 961

OY		962	TATGTCACTTGGAACTCATTGGAGTTCTGACCAAGCATCTGAAGAAAAGGTTCCA	1021
Db		962	ATAATGCATCTGGCAACTCATTTGGAGTTCTTGACCACAGCACTGGAAAAAGGTTCCA	1021
OY		1022	GTTTTCTGGGGATTAATTAACTCACAAGAAGGGAATTCAGCTGTAACTCATGCTACATTGAAAT	1081
Db		1022	GTTTTCTGGGGATTAATTAACTCACAAGAAGGGAATTCAGCTGTAACTCATGCTACATTGAAAT	1081
OY		1082	GCTGCATTTTATCCCTGAGTTTCAGGGATCGGAATCTCCCACCTCCAGAGACTTCMAATCATG	1141
Db		1082	GCTGCATTTTATCCCTGAGTTTCAGGGATCGGAATCTCCCACCTCCAGAGACTTCMAATCATG	1141
OY		1142	CGTGTGAAGCTCACTCGTCTTCATACATATTAGAATGTTTAGTGTGATGTCCTTGAGA	1201
Db		1142	CGTGTGAAGCTCACTCGTCTTCATACATATTAGAATGTTTAGTGTGATGTCCTTGAGA	1201
OY		1202	CATAGAGTTTGTGTATATCCGCAAAAGCTCTGAAACAGTAGAGGGGAATTAAGGGCTAA	1261
Db		1202	CATAGAGTTTGTGTATATCCGCAAAAGCTCTGAAACAGTAGAGGGGAATTAAGGGCTAA	1261
OY		1262	GATAGGAAGGTCGGCTCTTGTTGATATGTGSAANAATCTTAAAGAAAGTTGGTAGCTTTTCT	1321
Db		1262	GATAGGAAGGTCGGCTCTTGTTGATATGTGSAANAATCTTAAAGAAAGTTGGTAGCTTTTCT	1321
OY		1322	AGAGATTTCTGACCTTGAAAGATTAAAGAAAAAGCCAGGTGGCATATGCTTAAACGATAT	1381
Db		1322	AGAGATTTCTGACCTTGAAAGATTAAAGAAAAAGCCAGGTGGCATATGCTTAAACGATAT	1381
OY		1382	AACCTTGGGAACCTTAGGACGAGAGGGGTATAGTTCAAAGTTCAGCCAGGGCTATCTGTGTA	1441
Db		1382	AACCTTGGGAACCTTAGGACGAGAGGGGTATAGTTCAAAGTTCAGCCAGGGCTATCTGTGTA	1441
OY		1442	AGACTGTCTCMACATCCAAAGACCAAAATAAACAATPAGACAGACGAGAGGCTGAGATATGA	1501
Db		1442	AGACTGTCTCMACATCCAAAGACCAAAATAAACAATPAGACAGACGAGAGGCTGAGATATGA	1501
OY		1502	GAGCTGGACAGTGAAGTGATTTGTGTACAAACACGAGGAATCTATATTGATCTGAACCC	1561
Db		1502	GAGCTGGACAGTGAAGTGATTTGTGTGTACAAACACGAGGAATCTATATTGATCTGAACCC	1561
OY		1562	CCACAATGAAAAAGCTAGGCCCTGTATAGACATGCTGTATAGACTCAAGAAAGATGGAAGATTA	1621
Db		1562	CCACAATGAAAAAGCTAGGCCCTGTATAGACATGCTGTATAGACTCAAGAAAGATGGAAGATTA	1621
OY		1622	AGGCACAACAGATCCCCGGGGCTTGCCTGCAGTCAAGCTTAAGCTTAAGTGTCAATGCCAA	1681
Db		1622	AGGCACAACAGATCCCCGGGGCTTGCCTGCAGTCAAGCTTAAGCTTAAGTGTCAATGCCAA	1681
OY		1682	GTCCACAAGATCCCTGTCTCAMAGTAAAGATGAGTCAATATCTGAGGCCAATGTCATATGGG	1741
Db		1682	GTCCACAAGATCCCTGTCTCAMAGTAAAGATGAGTCAATATCTGAGGCCAATGTCATATGGG	1741
OY		1742	GATTGTCTCTCTCTCTCAGAAAGACATGCAATGACCCTGCACACACACACACACACAC	1801
Db		1742	GATTGTCTCTCTCTCTCAGAAAGACATGCAATGACCCTGCACACACACACACACACACAC	1801
OY		1802	ACACACACACACACACACACACACACACATAMATGAAGTTCTCTCTGTGCTGCTAACCC	1861
Db		1802	ACACACACACACACACACACACACACACATAMATGAAGTTCTCTCTGTGCTGCTAACCC	1861
OY		1862	TCTCATTAACATGTATCTCTTACAGAGACTCTCTGACCTCTGTTTAAAGCATGAGTGGGAG	1921
Db		1862	TCTCATTAACATGTATCTCTTACAGAGACTCTCTGACCTCTGTTTAAAGCATGAGTGGGAG	1921
OY		1922	CATGCGACAGACAGTCCAGTAATTTTATTCACAGACTCAAGAAAGCTGAGACAGAGCTGGA	1981
Db		1922	CATGCGACAGACAGTCCAGTAATTTTATTCACAGACTCAAGAAAGCTGAGACAGAGCTGGA	1981
OY		1982	GAGTTTCAAGAGACATGTGCCCCACATCTGTCCTTTACACAGAAAAAGTTTACC	2041
Db		1982	GAGTTTCAAGAGACATGTGCCCCACATCTGTCCTTTACACAGAAAAAGTTTACC	2041

Qy 2042 CGAAGCAGCTGCTGTCTGTAAAGAAACCTGCGAAGGCAAACTTTGACTGTG 2101  
 Db 2042 CGAAGCAGCTGCTGTCTGTAAAGAAACCTGCGAAGGCAAACTTTGACTGTG 2101  
 Qy 2102 TGGCTCAAGGGAACTGATCGAACAATTTCTCATTTCTGAGGAAATCGAGCTGTTTC 2161  
 Db 2102 TGGCTCAAGGGAACTGATCGAACAATTTCTCATTTCTGAGGAAATCGAGCTGTTTC 2161  
 Qy 2162 TGACAGAAAGCAACCGGTGACTGGGACATACGAAAGGAGAGCTTTGACAGCAATCTATA 2221  
 Db 2162 TGACAGAAAGCAACCGGTGACTGGGACATACGAAAGGAGAGCTTTGACAGCAATCTATA 2221  
 Qy 2222 TAGTCAGCAAAATATTTCTTGGAGAGACAGTGTCCACCAATTTGTTTCAAGCCGGTGG 2281  
 Db 2222 TAGTCAGCAAAATATTTCTTGGAGAGACAGTGTCCACCAATTTGTTTCAAGCCGGTGG 2281  
 Qy 2282 ACCCTCAGTTTATCTGGCTTACAGTGTGCTGCTGCCAGTCCCTTGATCTGTGCTGCC 2341  
 Db 2282 ACCCTCAGTTTATCTGGCTTACAGTGTGCTGCTGCCAGTCCCTTGATCTGTGCTGCC 2341  
 Qy 2342 ATCTATAACAGATCAATTAATTAATAGACCCGAGTGAATTAATTAAGTACAGCAAGAGT 2401  
 Db 2342 ATCTATAACAGATCAATTAATTAATAGACCCGAGTGAATTAATTAAGTACAGCAAGAGT 2401  
 Qy 2402 AGCTTTGTTCAAGATTTTTTGGCATTTGGGAGCAACTGTGTACATCAGAGACATCTGT 2461  
 Db 2402 AGCTTTGTTCAAGATTTTTTGGCATTTGGGAGCAACTGTGTACATCAGAGACATCTGT 2461  
 Qy 2462 TAGTGAGGACACCAAACTGTGTGTAACGTTTTTTCATGTATGAATTTTGTGTTAGGT 2521  
 Db 2462 TAGTGAGGACACCAAACTGTGTGTAACGTTTTTTCATGTATGAATTTTGTGTTAGGT 2521  
 Qy 2522 TGGCTTCTAGAGCTGTGGAGGCTGCTGCTTTCTTAGGTGGTATGAAAGGAGGACCATC 2581  
 Db 2522 TGGCTTCTAGAGCTGTGGAGGCTGCTGCTTTCTTAGGTGGTATGAAAGGAGGACCATC 2581  
 Qy 2582 TAACAAAATCCATTAGATTAAGATTAACAGCTCTCATGCAAGGAGGAAACTATCTCAATGTT 2641  
 Db 2582 TAACAAAATCCATTAGATTAAGATTAACAGCTCTCATGCAAGGAGGAAACTATCTCAATGTT 2641  
 Qy 2642 TTAAAGTAAATAAACTGTACTGGCAAGTCTTTGAGCATTAATAAAAAAAAAAAAAAAG 2701  
 Db 2642 TTAAAGTAAATAAACTGTACTGGCAAGTCTTTGAGCATTAATAAAAAAAAAAAAAAAG 2701  
 Qy 2702 GGGCGCGCG 2710  
 Db 2702 GGGCGCGCG 2710

RESULT 2  
 ARI68930 2710 bp DNA linear PAT 17-DEC-2001  
 LOCUS ARI68930 Sequence 8 from patent US 6288218.  
 DEFINITION ARI68930  
 ACCESSION ARI68930  
 VERSION ARI68930.1 GI:17905107  
 KEYWORDS  
 SOURCE unknown.  
 ORGANISM Unclassified.  
 REFERENCES 1 (bases 1 to 2710)  
 AUTHORS Levynson,D,Adam.  
 TITLE Compositions and methods for the treatment and diagnosis of immune disorders  
 JOURNAL Patent: US 6288218-A 8 11-SEP-2001;  
 FEATURES  
 SOURCE Location/Qualifiers  
 1..2710  
 /organism="unknown"  
 BASE COUNT 789 a 592 c 644 g 678 t 7 others  
 ORIGIN

Query Match 99.9%; Score 2706.6; DB 6; Length 2710;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GTGACCCACGCGTCCGGATTTTCCCTCCCAAGTACTATGTTTTCAGGTCTTACCTCA 61  
 Db 2 GTGACCCACGCGTCCGGATTTTCCCTCCCAAGTACTATGTTTTCAGGTCTTACCTCA 61  
 Qy 62 ACTGTGTCTGTCTGTCTGCACTACTTGTGCAAGGTCTTGAAGATGTTATAGG 121  
 Db 62 ACTGTGTCTGTCTGTCTGCACTACTTGTGCAAGGTCTTGAAGATGTTATAGG 121  
 Qy 122 TTGAGTTGGTAAAGGCTTATCTGCTCCGACCTTACCTTACCTTACCTGGAAC 181  
 Db 122 TTGAGTTGGTAAAGGCTTATCTGCTCCGACCTTACCTTACCTTACCTGGAAC 181  
 Qy 182 TTGTCCTTATGTGTGGGGCAAGGGATTTCTCTTGTGTCAGATGATCAATGAGTTC 241  
 Db 182 TTGTCCTTATGTGTGGGGCAAGGGATTTCTCTTGTGTCAGATGATCAATGAGTTC 241  
 Qy 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATCCAGGATACAGCTTAAAGGGCG 301  
 Db 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATCCAGGATACAGCTTAAAGGGCG 301  
 Qy 302 ATCTCAACAAAGGAGATGTGTCTGTATCATTAAGAAATGTACTTGTGATCAATGGGA 361  
 Db 302 ATCTCAACAAAGGAGATGTGTCTGTATCATTAAGAAATGTACTTGTGATCAATGGGA 361  
 Qy 362 CCTACTGTGCGAGATACAGTTCCCTGCTTATGAATGATTAATAATTAAGCTGAAT 421  
 Db 362 CCTACTGTGCGAGATACAGTTCCCTGCTTATGAATGATTAATAATTAAGCTGAAT 421  
 Qy 422 TAGACATCAAGGACGCAAGGTCACTCCAGCTCAGACTCCCATGAGGACTCTACTACAG 481  
 Db 422 TAGACATCAAGGACGCAAGGTCACTCCAGCTCAGACTCCCATGAGGACTCTACTACAG 481  
 Qy 482 CTTCTCCAAAGACCTTAAACACGAGAGAAATGTGTTAGAGACACACACACTGTGATCC 541  
 Db 482 CTTCTCCAAAGACCTTAAACACGAGAGAAATGTGTTAGAGACACACACACTGTGATCC 541  
 Qy 542 TCCATTAATTAACATGGAACAAATAATTTCCATGAGGCTGATGAATTAAGACTGTGAG 601  
 Db 542 TCCATTAATTAACATGGAACAAATAATTTCCATGAGGCTGATGAATTAAGACTGTGAG 601  
 Qy 602 AAACGATCAGAACTGTATCCATTCGATGAGAGTCTGTCTGTGAGTGAACCTGTGAC 661  
 Db 602 AAACGATCAGAACTGTATCCATTCGATGAGAGTCTGTCTGTGAGTGAACCTGTGAC 661  
 Qy 662 TTATCATTTGGTGTCTTAATCTTAAATGTATTCCTGTAAAGAAAGAAATTCAGATT 721  
 Db 662 TTATCATTTGGTGTCTTAATCTTAAATGTATTCCTGTAAAGAAAGAAATTCAGATT 721  
 Qy 722 TGAGCTTATTAACCTGGCCAACTTGCTCCAGAGAGGTTGGCAATGACAGAGACATCA 781  
 Db 722 TGAGCTTATTAACCTGGCCAACTTGCTCCAGAGAGGTTGGCAATGACAGAGACATCA 781  
 Qy 782 GGATTCGCTGTAGAGAAATAATCTTACACCATCGAGAGAAAGTATGATGAGAGAAAT 841  
 Db 782 GGATTCGCTGTAGAGAAATAATCTTACACCATCGAGAGAAAGTATGATGAGAGAAAT 841  
 Qy 842 CAAATGAGTACTACTGTACTACGTCACACGACGACGACCTCTGAGCTGCC 901  
 Db 842 CAAATGAGTACTACTGTACTACGTCACACGACGACGACCTCTGAGCTGCC 901  
 Qy 902 ACTTTTAAAGGCTGCGCTTCAATTTCTGACTTTGATTTCCCTTGTGAGAAACTATGTG 961  
 Db 902 ACTTTTAAAGGCTGCGCTTCAATTTCTGACTTTGATTTCCCTTGTGAGAAACTATGTG 961  
 Qy 962 ATATGTACTTGGCAACCTCATTTGAGAGTTCTGACACAGCACTGAGAGAAAGATTCCA 1021  
 Db 962 ATATGTACTTGGCAACCTCATTTGAGAGTTCTGACACAGCACTGAGAGAAAGATTCCA 1021  
 Qy 1022 GTTTTCTGGGATTAATTAATCAACAGGAGTTGACTGTACTATCTCATTTGAAT 1081  
 Db 1022 GTTTTCTGGGATTAATTAATCAACAGGAGTTGACTGTACTATCTCATTTGAAT 1081

QY	1082	GCTCATTATTAATCCCTGAGTTTACAGGAATCGGATCTCCCACTCCAGAACATTCATATG	11411
Db	1082	GCTCATTATTAATCCCTGAGTTTACAGGAATCGGATCTCCCACTCCAGAACATTCATATG	11411
QY	1142	CGTGTGAAGCTCACTCGTGCCTTTCATATCATTAGAATGTTAGTGTATCTTTGAGA	1201
Db	1142	CGTGTGAAGCTCACTCGTGCCTTTCATATCATTAGAATGTTAGTGTATCTTTGAGA	1201
QY	1202	CATAGAGTTTGTGTATATCCGCAAACTCTGTAAACAGGTATAGGGGAAATAAGGCTTAA	1261
Db	1202	CATAGAGTTTGTGTATATCCGCAAACTCTGTAAACAGGTATAGGGGAAATAAGGCTTAA	1261
QY	1262	GATTAGGAAGTCCGCTCTTGTGTATGTGGAAAATCTTAAAGAAATGTGATAGCTTTTCT	1321
Db	1262	GATTAGGAAGTCCGCTCTTGTGTATGTGGAAAATCTTAAAGAAATGTGATAGCTTTTCT	1321
QY	1322	AGAGATTTCTGACTTGTGAAGATTAAGAAAAAGCCAGGTGGCATATGCTTTAAACGATAT	1381
Db	1322	AGAGATTTCTGACTTGTGAAGATTAAGAAAAAGCCAGGTGGCATATGCTTTAAACGATAT	1381
QY	1382	AACTTGGGAACCTTAGGCAAGAGGCTGTAAAGTTCAAGTCAAGCCAGGCTATGCTGTA	14411
Db	1382	AACTTGGGAACCTTAGGCAAGAGGCTGTAAAGTTCAAGTCAAGCCAGGCTATGCTGTA	14411
QY	1442	AGACTGTCTCMACATCCCAAAAGCGAAATTAACATAGAGACACAGAGGCTGTGAGATGA	1501
Db	1442	AGACTGTCTCMACATCCCAAAAGCGAAATTAACATAGAGACACAGAGGCTGTGAGATGA	1501
QY	1502	GGCTTCGACAGTGAAGTGCATTTGTATTAAGACACAGAAATCTATATTGATCTGATACC	1561
Db	1502	GGCTTCGACAGTGAAGTGCATTTGTATTAAGACACAGAAATCTATATTGATCTGATACC	1561
QY	1562	CCACATGAAAAAGCTAAGGCTGTGTAGACATGCTTGTAGACTCAAGAGATGAGAGTTAA	1621
Db	1562	CCACATGAAAAAGCTAAGGCTGTGTAGACATGCTTGTAGACTCAAGAGATGAGAGTTAA	1621
QY	1622	AGGCAACAAGATCCCGGGGCTTCCGAGCAAGTCAAGCTTAAAGTGGCTGAGTTCCAA	1681
Db	1622	AGGCAACAAGATCCCGGGGCTTCCGAGCAAGTCAAGCTTAAAGTGGCTGAGTTCCAA	1681
QY	1682	GTCACAAGAGTCCCTGTCTCMAAGTAAGATGATGTGAATGCTGGCCATGTCCATGGG	1741
Db	1682	GTCACAAGAGTCCCTGTCTCMAAGTAAGATGATGTGAATGCTGGCCATGTCCATGGG	1741
QY	1742	GGTGTGCTCTCTCTCAGAAAGACATGCAATGACCCTGACACACACACACACACAC	1801
Db	1742	GGTGTGCTCTCTCTCAGAAAGACATGCAATGACCCTGACACACACACACACACAC	1801
QY	1802	ACACACACACACACACACACACACACACATGMAATGAAGTGTCTTGTGTGCTGTACC	1861
Db	1802	ACACACACACACACACACACACACACACATGMAATGAAGTGTCTTGTGTGCTGTACC	1861
QY	1862	TCTCATTAACATGTAATCTCTCAAGAGCTCTCTGCTCTGTTTAAACATGAGTGGAG	1921
Db	1862	TCTCATTAACATGTAATCTCTCAAGAGCTCTCTGCTCTGTTTAAACATGAGTGGAG	1921
QY	1922	CATGCAAGACAGTCCAGTAATTTATTCACAGCACTCAAGAGCTGTGAGACAGAGCTGGA	1981
Db	1922	CATGCAAGACAGTCCAGTAATTTATTCACAGCACTCAAGAGCTGTGAGACAGAGCTGGA	1981
QY	1982	GAGTTTCAGAGACATGTGTCCCAACCTGCAGACTCTTCAACACMAAAAAAGTTATCC	2041
Db	1982	GAGTTTCAGAGACATGTGTCCCAACCTGCAGACTCTTCAACACMAAAAAAGTTATCC	2041
QY	2042	CGCAAGACGCTGCTGTGTATAAGAAACCTCGCAAAAGGCAAACTTTGACTGTGTG	2101
Db	2042	CGCAAGACGCTGCTGTGTATAAGAAACCTCGCAAAAGGCAAACTTTGACTGTGTG	2101
QY	2102	TGCTCAAGGGAACTGACTCAGACAACTTCTCATTTCTGTGAGAAATCTGAGCTGTTTC	2161
Db	2102	TGCTCAAGGGAACTGACTCAGACAACTTCTCATTTCTGTGAGAAATCTGAGCTGTTTC	2161
QY	2162	TGACAGAAAGAACCGGTGACTGAGACATACAAAGGACAGACTTTCGACATCTATA	2221

Db	2162	TGACAGAAAGAACCAACCGGTGACTGGGACATACGAAAGGCAGAGCTCTTGACGAAATCTATA	2221
OY	2222	TAGTCAGCAAAATATCTTTGGGAGAGACAGTCGTCAACCAATTGATTTCCAGCCGGTGG	2281
Db	2222	TAGTCACACAAATATCTTTGGGAGAGACAGTCGTCAACCAATTGATTTCCAGCCGGTGG	2281
OY	2282	ACCTCAGTTTCATGTGGCTTACAGCTGCTGCCAGTGCCCTTGATCTGTGCTGGCTCCC	2341
Db	2282	ACCTCAGTTTCATGTGGCTTACAGCTGCTGCCAGTGCCCTTGATCTGTGCTGGCTCCC	2341
OY	2342	ATCTATAACGAATCAAAATTAAATAGACCCCGAGTGAATAATTAACTAGTGACAGAAAGT	2401
Db	2342	ATCTATAACGAATCAAAATTAAATAGACCCCGAGTGAATAATTAACTAGTGACAGAAAGT	2401
OY	2402	AGCTTTTTCAAAAGATTTTTTGATTTGGGAGAGCACTGTGTACATAGAGGACATCTGT	2461
Db	2402	AGCTTTTTCAAAAGATTTTTTGATTTGGGAGAGCACTGTGTACATAGAGGACATCTGT	2461
OY	2462	TAGTGAGAGACACCAAAACCTGTGGTACCGTTTTTTCATGTATGAATTTGTGTTTAGT	2521
Db	2462	TAGTGAGAGACACCAAAACCTGTGGTACCGTTTTTTCATGTATGAATTTGTGTTTAGT	2521
OY	2522	TGCTTTTAGCTAGCTGTGTGAGGTCTGTGGCTTTTCTTAGGTGGGTATGGAAGGAGACATC	2581
Db	2522	TGCTTTTAGCTAGCTGTGTGAGGTCTGTGGCTTTTCTTAGGTGGGTATGGAAGGAGACATC	2581
OY	2582	TAAACAAATCCATTAGAGATPAACAGCTCTCATGCGAGAAAGGAAAACTAATCTCAAAATGTT	2641
Db	2582	TAAACAAATCCATTAGAGATPAACAGCTCTCATGCGAGAAAGGAAAACTAATCTCAAAATGTT	2641
OY	2642	TTAAAGTAAATAAACCTGTACTGGCAAGATCTTTAGCATTAACCAATCAAAAAAAAAAAG	2701
Db	2642	TTAAAGTAAATAAACCTGTACTGGCAAGATCTTTAGCATTAACCAATCAAAAAAAAAAAG	2701
OY	2702	GGCGGCGGC 2710	
Db	2702	GGCGGCGGC 2710	

RESULT 3					PAT 20-DEC-2002
AR232678					
LOCUS	AR232678	2710 bp	DNA	linear	
DEFINITION	Sequence 8 from patent US 6455665.				
ACCESSION	AR232678				
VERSION	AR232678.1	GI:27274955			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 2710)				
TITLE	Levinson, D.A. Compositions and methods for the treatment and diagnosis of immune disorders				
JOURNAL FEATURES	Patent: US 6455665-A 8 24-SEP-2002; Location/Qualifiers 1..2710 . /organism="unknown"				
BASE COUNT	789 a . 592 c 644 g 678 t			7 others	
ORIGIN					

Query Match	Similarity	Score	DB	Length
99.9%	100.0%	2706.6	6	2710
Beet Local	Similarity	Pred. No. 0		
Matches	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0

  

Query	Sequence	Length
QY	2 GTGCAGCCAGCGCGGATTTCCCTCCAGACTCATGTTTCAAGTCTTACCTCA	61
DB	2 GTGCAGCCAGCGCGGATTTCCCTCCAGACTCATGTTTCAAGTCTTACCTCA	61
QY	62 ACTGTGTCCTCTCTGTGCACTACTTGTGAAGGTCTTGAAGATGTTATAAG	121
DB	62 ACTGTGTCCTCTCTGTGCACTACTTGTGAAGGTCTTGAAGATGTTATAAG	121

122 TTGAGTTGGTAAAAATGCTATCTGCCCCGCTGAGTTACACTCTACTCATCTGGGACAC 181  
122 TTGAGTTGGTAAAAATGCTATCTGCCCCGCTGAGTTACACTCTACTCATCTGGGACAC 181  
182 TTGTCCTATATGCTGGGGCAAGGGATTTCTGCTCTTGGTCAAGTGTACCAATGAGTTGC 241  
182 TTGTCCTATATGCTGGGGCAAGGGATTTCTGCTCTTGGTCAAGTGTACCAATGAGTTGC 241  
242 TCAGAACTGATGAAAGAAATGACATATCCAGAAATCCAGAGATACCAAGCTAAAGGGCG 301  
242 TCAGAACTGATGAAAGAAATGACATATCCAGAAATCCAGAGATACCAAGCTAAAGGGCG 301  
302 ATCTCAACAAAGAGATGTTCTCTGATCATTAAGAAATGTGACTCTGGATGCATGGGA 361  
302 ATCTCAACAAAGAGATGTTCTCTGATCATTAAGAAATGTGACTCTGGATGCATGGGA 361  
362 CTTACTGCTGCAAGATACAGTTCCCTGCTCTTATGATGATTAAGAAATTAAGACTGAAT 421  
362 CTTACTGCTGCAAGATACAGTTCCCTGCTCTTATGATGATTAAGAAATTAAGACTGAAT 421  
422 TAGACATTAAGAGCCCAAGGCTCACTCCAGCTCAGACTGCCATGGGAGCTTACTAGC 481  
422 TAGACATTAAGAGCCCAAGGCTCACTCCAGCTCAGACTGCCATGGGAGCTTACTAGC 481  
482 CTTCTCCAAGAACCTTAACACAGAGAGAAATGTTTCAAGACACAGACACTGGTGACC 541  
482 CTTCTCCAAGAACCTTAACACAGAGAGAAATGTTTCAAGACACAGACACTGGTGACC 541  
542 TCCATTAATTAACATGGAACAAAATTTTCCATGAGGCTGATGAATTAAGACCTCTGAG 601  
542 TCCATTAATTAACATGGAACAAAATTTTCCATGAGGCTGATGAATTAAGACCTCTGAG 601  
602 AAAAGATAGAACTGCTATCCACATTTGAGAGTGGAGTCTCTGGGTTGACCTCTGGAC 661  
602 AAAAGATAGAACTGCTATCCACATTTGAGAGTGGAGTCTCTGGGTTGACCTCTGGAC 661  
662 TTATCATTTGTTCTTAATCTTAATGATGATTTCTCTGAGAAAGAAAGATTAATGATG 721  
662 TTATCATTTGTTCTTAATCTTAATGATGATTTCTCTGAGAAAGAAAGATTAATGATG 721  
722 TGAAGCTTATTAACATGAGCCAACTTGGCTTCCAGAGAGGTTGGCAATGACAGAGCAGTCA 781  
722 TGAAGCTTATTAACATGAGCCAACTTGGCTTCCAGAGAGGTTGGCAATGACAGAGCAGTCA 781  
782 GGAATTCGCTGAGAGAAATATCTACACCATGAGAGAAAGATTAATGAGTGGAGATT 841  
782 GGAATTCGCTGAGAGAAATATCTACACCATGAGAGAAAGATTAATGAGTGGAGATT 841  
842 CAAATGAGTACTACTGCTAGCTCAACAGCAGCAGCAGCCATCTGACCCGCTCTGAGCTGC 901  
842 CAAATGAGTACTACTGCTAGCTCAACAGCAGCAGCAGCCATCTGACCCGCTCTGAGCTGC 901  
902 ACTTTTAAAGGCTGCTTCAATTTCTGAATTTGATTTTCCCTTTTGTGAAAACTATGTG 961  
902 ACTTTTAAAGGCTGCTTCAATTTCTGAATTTGATTTTCCCTTTTGTGAAAACTATGTG 961  
962 ACTTTTAAAGGCTGCTTCAATTTCTGAATTTGATTTTCCCTTTTGTGAAAACTATGTG 961  
962 ACTTTTAAAGGCTGCTTCAATTTCTGAATTTGATTTTCCCTTTTGTGAAAACTATGTG 961  
962 ATATGCTCACTTGGGCAACCTCATTTGAGAGTGTGACCAACAGCAGCTGAGAAAAAGATTCCA 1021  
962 ATATGCTCACTTGGGCAACCTCATTTGAGAGTGTGACCAACAGCAGCTGAGAAAAAGATTCCA 1021  
1022 GTTTTCTGGGATTAATTAATCAACAGGGGATTCGACTGTAACTCATCTCATTTGAAT 1081  
1022 GTTTTCTGGGATTAATTAATCAACAGGGGATTCGACTGTAACTCATCTCATTTGAAT 1081  
1082 GCTCCATTTTATCCCTGAGTTTCAAGGATTCGATCTCCCACTCCAGAGACTTCAATCATG 1141  
1082 GCTCCATTTTATCCCTGAGTTTCAAGGATTCGATCTCCCACTCCAGAGACTTCAATCATG 1141  
1142 GGTGTTGAAGCTCACTGCTGCTTCAATTAAGAAATGTTAGTGTATGCTTTTGA 1201  
1142 GGTGTTGAAGCTCACTGCTGCTTCAATTAAGAAATGTTAGTGTATGCTTTTGA 1201  
1202 CATAGAGTTTGTGATATCCGCAAGCTCCTGAAACAGTAGGGGAAATAAGGGCTAA 1261

1202 CATAGAGTTTGTGATATCCGCAAGCTCCTGAACAGTAGGGGAAATAAGGGCTAA 1261  
1262 GATAGAAAGTGGCTCTTTGTTGATGTTGAAAAATCTTAAGAAATGTTGATGCTTTTCT 1321  
1262 GATAGAAAGTGGCTCTTTGTTGATGTTGAAAAATCTTAAGAAATGTTGATGCTTTTCT 1321  
1322 AGAGATTTCTGACCTTGAAAGATTAAGAAAGAGCAGGTGCAATATGCTTAACAGATAT 1381  
1322 AGAGATTTCTGACCTTGAAAGATTAAGAAAGAGCAGGTGCAATATGCTTAACAGATAT 1381  
1382 AACTTGGGAACTTAAAGCAGAGGTTGATTAAGTTCAAGGTGAGCCAGGCTATGCTGTA 1441  
1382 AACTTGGGAACTTAAAGCAGAGGTTGATTAAGTTCAAGGTGAGCCAGGCTATGCTGTA 1441  
1442 AGACTGCTCAGCATCCAAAGACAAATTAACATAGAGACAGCAGAGGCTGAGATGA 1501  
1442 AGACTGCTCAGCATCCAAAGACAAATTAACATAGAGACAGCAGAGGCTGAGATGA 1501  
1502 GGTCTGGAAGTGGAGTGCATTTGTTGTAACAGCAGAGAACTATATTTGATCTGAGAC 1561  
1502 GGTCTGGAAGTGGAGTGCATTTGTTGTAACAGCAGAGAACTATATTTGATCTGAGAC 1561  
1562 GGTCTGGAAGTGGAGTGCATTTGTTGTAACAGCAGAGAACTATATTTGATCTGAGAC 1561  
1562 GGTCTGGAAGTGGAGTGCATTTGTTGTAACAGCAGAGAACTATATTTGATCTGAGAC 1561  
1622 AGGCAACAAGATCCCGGGGCTTGGCTGAGTCACTTACCTTACCTGAGTTCCAA 1681  
1622 AGGCAACAAGATCCCGGGGCTTGGCTGAGTCACTTACCTTACCTGAGTTCCAA 1681  
1622 AGGCAACAAGATCCCGGGGCTTGGCTGAGTCACTTACCTTACCTGAGTTCCAA 1681  
1682 GTCCAAAGATCCCTGCTCAGAGTAAAGTGTGATGATGAGTGTGAGTGTGAGTGTGAGT 1741  
1682 GTCCAAAGATCCCTGCTCAGAGTAAAGTGTGATGATGAGTGTGAGTGTGAGTGTGAGT 1741  
1742 GGTGCTCTCTCTCTCAGAAAGACATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1801  
1742 GGTGCTCTCTCTCTCAGAAAGACATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1801  
1802 ACAACACACACACACACACACACACACACATGAMATGAAAGTTCTCTGCTGCTGCTAC 1861  
1802 ACAACACACACACACACACACACACACACATGAMATGAAAGTTCTCTGCTGCTGCTAC 1861  
1862 TCTCTATTAACATGATATCTCTACAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1921  
1862 TCTCTATTAACATGATATCTCTACAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1921  
1922 CATGACAGAGAGTCCAGTAAATTTATTCAGACACTCAGAAAGCTGAGCAGAGGCTGGA 1981  
1922 CATGACAGAGAGTCCAGTAAATTTATTCAGACACTCAGAAAGCTGAGCAGAGGCTGGA 1981  
1982 GAGTTCAAGAGCACTGTGCCCAACACTGCGAGACTCTTCTTAACAAAGAAAGGTTACC 2041  
1982 GAGTTCAAGAGCACTGTGCCCAACACTGCGAGACTCTTCTTAACAAAGAAAGGTTACC 2041  
2042 CGCAAGAGCCTGCTGCTGTTGTAAGAAAGAAACCTGCGAAAGCAAACTTTGACTGTTG 2101  
2042 CGCAAGAGCCTGCTGCTGTTGTAAGAAAGAAACCTGCGAAAGCAAACTTTGACTGTTG 2101  
2102 TGCTCAAGGGGAACTGACTCAGACAACTTCCCAATCTGAGAGAACTGAGAGCTGTTTC 2161  
2102 TGCTCAAGGGGAACTGACTCAGACAACTTCCCAATCTGAGAGAACTGAGAGCTGTTTC 2161  
2162 TGAAGAAAGCAACCGGTGACTGAGCAATGAGAGGAGAGCTTTGACAGCAATCTATA 2221  
2162 TGAAGAAAGCAACCGGTGACTGAGCAATGAGAGGAGAGCTTTGACAGCAATCTATA 2221  
2222 TAGTCAGAAATTAATCTTTGAGAGACAGTGTGCAACAAATTAATTTTCAAGCCGTTG 2281  
2222 TAGTCAGAAATTAATCTTTGAGAGACAGTGTGCAACAAATTAATTTTCAAGCCGTTG 2281  
2282 AACTCAGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2341  
2282 AACTCAGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2341

Db 2282 ACCTGAGTTTCATCTGCTTACAGCTGCTGCCAGTCCCTTGATCTGTGCTGCTCC 2341

Qy 2342 ATCTATACAGAAATCAATTAATTAAGACCCGAGTGAATAATTAAGTACAGAAAGT 2401

Db 2342 ATCTATACAGAAATCAATTAATTAAGACCCGAGTGAATAATTAAGTACAGAAAGT 2401

Qy 2402 ACCTTGTTCAGAAATTTTGTGATTTGGGAGCACTGTGTATCATACAGAGCATCTGT 2461

Db 2402 ACCTTGTTCAGAAATTTTGTGATTTGGGAGCACTGTGTATCATACAGAGCATCTGT 2461

Qy 2462 TAGTGAAGACACCAAAACCTGTGTACCCGTTTTCATGATGATATTTTGTGTAGGT 2521

Db 2462 TAGTGAAGACACCAAAACCTGTGTACCCGTTTTCATGATGATATTTTGTGTAGGT 2521

Qy 2522 TGGCTTACCTAGCTGTGAGAGTCTGTGCTTTCTTAAAGGTATGAAAGGAGACATC 2581

Db 2522 TGGCTTACCTAGCTGTGAGAGTCTGTGCTTTCTTAAAGGTATGAAAGGAGACATC 2581

Qy 2582 TAAACAAATCCATTAGATTAACAGCTTCACAGCAAGGAAACTTAATCTCAATGTT 2641

Db 2582 TAAACAAATCCATTAGATTAACAGCTTCACAGCAAGGAAACTTAATCTCAATGTT 2641

Qy 2642 TTAAGATAATTAATCTGTACCTGTGCAAGTACTTTGAGCATTAATAAAAAAAAAAAG 2701

Db 2642 TTAAGATAATTAATCTGTACCTGTGCAAGTACTTTGAGCATTAATAAAAAAAAAAAG 2701

Qy 2702 GGGGGGCGC 2710

Db 2702 GGGGGGCGC 2710

RESULT 4

189415

LOCUS 189415 2710 bp DNA linear PAT 10-AUG-1998

DEFINITION Sequence 8 from patent US 5721351.

ACCESSION 189415

VERSION 189415.1 GI:3409355

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 2710)

AUTHORS Levinson,D,Adam.

TITLE Compositions and methods for the treatment and diagnosis of immune disorders

JOURNAL Patent: US 5721351-A 8 24-FEB-1998;

FEATURES

source Location/Qualifiers

BASE COUNT 789 a 592 c 644 g 678 t 7 others

ORIGIN

Query Match 99.9%; Score 2706.6; DB 6; Length 2710;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 242 TCAGAACTGATATAAGAAATGTGACATATCAAAATCCAGAGATACAGCTTAAGGGCG 301

Qy 302 ATCTCAACAAAGAGATGTGTCTGATCATTAAGAAATGTGATCTGATGACCAATGGGA 361

Db 302 ATCTCAACAAAGAGATGTGTCTGATCATTAAGAAATGTGATCTGATGACCAATGGGA 361

Qy 362 CCTACTGCTGACAGATACAGTTCCTGGTCTTATGATGATTAATAAATTAGAACTGAAT 421

Db 362 CCTACTGCTGACAGATACAGTTCCTGGTCTTATGATGATTAATAAATTAGAACTGAAT 421

Qy 422 TAGACATCAAGACAGCAAGGTCATCCAGCTCAGACTGGCCGATGGGAGCTACTACAG 481

Db 422 TAGACATCAAGACAGCAAGGTCATCCAGCTCAGACTGGCCGATGGGAGCTACTACAG 481

Qy 482 CTCTCCAGAAACCCCTAACCAAGAGAAATGTTCAGAGACACAGCACTGTGACCC 541

Db 482 CTCTCCAGAAACCCCTAACCAAGAGAAATGTTCAGAGACACAGCACTGTGACCC 541

Qy 542 TCCATTAATTAATGAAACAAATTTTCCATGTGGCTGATGAATAATTAAGACTGTGAG 601

Db 542 TCCATTAATTAATGAAACAAATTTTCCATGTGGCTGATGAATAATTAAGACTGTGAG 601

Qy 602 AAACGATCAGAACTGCTATCCATTTGAGAGTGGAGTCTGTGGGTTGACCTGGGAC 661

Db 602 AAACGATCAGAACTGCTATCCATTTGAGAGTGGAGTCTGTGGGTTGACCTGGGAC 661

Qy 662 TTATCATTTGCTCTTAATCTTAATATGATATCTCTGTAAGAAAGAAATGATGAGTT 721

Db 662 TTATCATTTGCTCTTAATCTTAATATGATATCTCTGTAAGAAAGAAATGATGAGTT 721

Qy 722 TGAGCTTATTAACCTGGCAACCTTGCTCCAGAGGGTTGGCAATGCAAGAGCATCA 781

Db 722 TGAGCTTATTAACCTGGCAACCTTGCTCCAGAGGGTTGGCAATGCAAGAGCATCA 781

Qy 782 GGAATTCGCTGAGGAAATATCTACCATGAGAGAGAAAGTATATGAGGAGAAAT 841

Db 782 GGAATTCGCTGAGGAAATATCTACCATGAGAGAGAAAGTATATGAGGAGAAAT 841

Qy 842 CAAATGAGTACTACTGCTACAGTCAACAGCAGCAGCACTCTGACCGCTCTGACCTGCC 901

Db 842 CAAATGAGTACTACTGCTACAGTCAACAGCAGCAGCACTCTGACCGCTCTGACCTGCC 901

Qy 902 ACTTTTAAAGGCTGGCTTCAATTTCTGACTTTGGATTTTCCCTTTTGTGAAAACTATG 961

Db 902 ACTTTTAAAGGCTGGCTTCAATTTCTGACTTTGGATTTTCCCTTTTGTGAAAACTATG 961

Qy 962 ATATGCACTTGGCAACCTCATTTGAGGTTCTGACCAAGCCACTGAGAAAAAGATTCCA 1021

Db 962 ATATGCACTTGGCAACCTCATTTGAGGTTCTGACCAAGCCACTGAGAAAAAGATTCCA 1021

Qy 1022 GTTTTCTGGGATAATTAATCAACAAAGGATTCGATGTAATCTAGCTACATGTAAT 1081

Db 1022 GTTTTCTGGGATAATTAATCAACAAAGGATTCGATGTAATCTAGCTACATGTAAT 1081

Qy 1082 GCTTCATTTTATCCCTGAGTTTCAAGGATCGGATCTCCCACTCCAGAGCTTCAATCAAG 1141

Db 1082 GCTTCATTTTATCCCTGAGTTTCAAGGATCGGATCTCCCACTCCAGAGCTTCAATCAAG 1141

Qy 1142 CGTGTGAAGCTCAGCTCGGCTTTCATATCAATTAAGAAATGTTAGTGTCTTTGAGA 1201

Db 1142 CGTGTGAAGCTCAGCTCGGCTTTCATATCAATTAAGAAATGTTAGTGTGTCTTTGAGA 1201

Qy 1202 CATAGAGTTTGTGATATCCGAAAGCTCTCGAAACAGTGAAGGGAATAAAGGGCTAA 1261

Db 1202 CATAGAGTTTGTGATATCCGAAAGCTCTCGAAACAGTGAAGGGAATAAAGGGCTAA 1261

Qy 1262 GATAGAGGTCGGCTTTTGTGATTTTGAATAATCTTAAGAAAGTTGTGCTTTTCT 1321

Db 1262 GATAGAGGTCGGCTTTTGTGATTTTGAATAATCTTAAGAAAGTTGTGCTTTTCT 1321

Qy 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCAGGGGCTGATGCTTAACAGATAT 1381

Db 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCAGGGGCTGATGCTTAACAGATAT 1381

Db 1322 AGAGATTTCGACCTTGAAAGATTAAAGAAAAAGCCAGGTGGCATATGCTTAACACGATAT 1381  
 Qy 1382 AACTTGGGAACCTTAGGAGAGGGGTATAGTTCAAGSTACGACAGGGCTATCTGTA 1441  
 Db 1382 AACTGGGAACCTTAGGAGAGGGGTATAGTTCAAGSTACGACAGGGCTATCTGTA 1441  
 Qy 1442 AGACTGTCTCMCATCCAAAGCAAAATAAACAATAGACAGAGAGAGGCTGAGATGA 1501  
 Db 1442 AGACTGTCTCMCATCCAAAGCAAAATAAACAATAGACAGAGAGAGGCTGAGATGA 1501  
 Qy 1502 GGGTCGACAGTGAAGTGCATTTGTGTACAGACGAGGAATCTATTTGATCGTAGAC 1561  
 Db 1502 GGGTCGACAGTGAAGTGCATTTGTGTACAGACGAGGAATCTATTTGATCGTAGAC 1561  
 Qy 1562 CCAATGAAAAAGCTAGGCGCTGTAGAGCATGCTTGTATGACTCAAGAGATGAGAGGTA 1621  
 Db 1562 CCAATGAAAAAGCTAGGCGCTGTAGAGCATGCTTGTATGACTCAAGAGATGAGAGGTA 1621  
 Qy 1622 AGGCAACAAGATCCCGGGGCTTGCGTGCATGAGTCAAGCTTAGAGTGTAGTCCAA 1681  
 Db 1622 AGGCAACAAGATCCCGGGGCTTGCGTGCATGAGTCAAGCTTAGAGTGTAGTCCAA 1681  
 Qy 1682 GTCCACAAGAGTCCCTGTCTCAMAAGTAAGATGAGTGAATCTGGCGCATGTCCATGG 1741  
 Db 1682 GTCCACAAGAGTCCCTGTCTCAMAAGTAAGATGAGTGAATCTGGCGCATGTCCATGG 1741  
 Qy 1742 GGTGTCTCTCCCTCTGAGAGACATGACATGACCTTCCACACACACACACACACAC 1801  
 Db 1742 GGTGTCTCTCTCTGAGAGACATGACATGACCTTCCACACACACACACACACAC 1801  
 Qy 1802 ACACACACACACACACACACACACACATGAMATGAGTCTCTGCTGCTCTCTCTCTCT 1861  
 Db 1802 ACACACACACACACACACACACACACATGAMATGAGTCTCTGCTGCTCTCTCTCTCT 1861  
 Qy 1862 TCTCTATTAACATGATCTCTCAAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
 Db 1862 TCTCTATTAACATGATCTCTCAAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
 Qy 1922 CATGCGACAGAGCTCAAGTATTTATTCACACACTCAAGAGCTGAGACAGAGGTGGA 1981  
 Db 1922 CATGCGACAGAGCTCAAGTATTTATTCACACACTCAAGAGCTGAGACAGAGGTGGA 1981  
 Qy 1982 GAGTTCAAGAGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2041  
 Db 1982 GAGTTCAAGAGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2041  
 Qy 2042 CGCAAGCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2101  
 Db 2042 CGCAAGCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2101  
 Qy 2102 TGCTCAAGGGGAACTGATCTGACAACTTTCTTCATCTTCTGAGAGAACTGAGAGCTGTTT 2161  
 Db 2102 TGCTCAAGGGGAACTGATCTGACAACTTTCTTCATCTTCTGAGAGAACTGAGAGCTGTTT 2161  
 Qy 2162 TGACAGAAGAACCAACCGGTGACTGGGACATACGAAGCAGAGCTCTTGACAGAACTCTATA 2221  
 Db 2162 TGACAGAAGAACCAACCGGTGACTGGGACATACGAAGCAGAGCTCTTGACAGAACTCTATA 2221  
 Qy 2222 TAGTCAGAAAAATATTTCTTGGGAGAGAGTGTACCAAAATTTGATTTCAAGCCGGTGG 2281  
 Db 2222 TAGTCAGAAAAATATTTCTTGGGAGAGAGTGTACCAAAATTTGATTTCAAGCCGGTGG 2281  
 Qy 2282 AACTTCAAGTTTCACTGGCTTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2341  
 Db 2282 AACTTCAAGTTTCACTGGCTTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2341  
 Qy 2342 ATCTATAACAGATCAAAATTAATAGACCCGAGTGAATAATTAAGTGAAGCAAAAGT 2401  
 Db 2342 ATCTATAACAGATCAAAATTAATAGACCCGAGTGAATAATTAAGTGAAGCAAAAGT 2401  
 Qy 2402 AGCTTTGTTCAAAAGATTTTGTGCAATGGGAGCAACTGTGTACATCAGAGACATCTGT 2461  
 Db 2402 AGCTTTGTTCAAAAGATTTTGTGCAATGGGAGCAACTGTGTACATCAGAGACATCTGT 2461

Qy 2462 TAGTGAAGACACCAAAACCTGTGTACGTTTTTTCATGATGAATTTTGTGTTAGGT 2521  
 Db 2462 TAGTGAAGACACCAAAACCTGTGTACGTTTTTTCATGATGAATTTTGTGTTAGGT 2521  
 Qy 2522 TGCTTTAGTAGTGTGTGAGAGTCTGCTGCTTCTTCTTGAAGGAGTGAAGAGGAGACCATC 2581  
 Db 2522 TGCTTTAGTAGTGTGTGAGAGTCTGCTGCTTCTTCTTGAAGGAGTGAAGAGGAGACCATC 2581  
 Qy 2582 TTAACAAATCCATTAGATTAACAGCTCTCATGCAAGAGGAAACCTAATCTCAATTT 2641  
 Db 2582 TTAACAAATCCATTAGATTAACAGCTCTCATGCAAGAGGAAACCTAATCTCAATTT 2641  
 Qy 2642 TTAAGTAATTAACCTGACGCAAGTACTTTGACATTAATAAAAAAAAAAAAAAG 2701  
 Db 2642 TTAAGTAATTAACCTGACGCAAGTACTTTGACATTAATAAAAAAAAAAAAAAG 2701  
 Qy 2702 GCGCGCCGC 2710  
 Db 2702 GCGCGCCGC 2710  
 RESULT 5  
 AF450241  
 LOCUS  
 DEFINITION Mus musculus T cell immunoglobulin mucin-3 mRNA, complete cds.  
 ACCESSION AF450241  
 VERSION AF450241.1 GI:18182530  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS  
 1 (bases 1 to 2725)  
 Monney, L., Sabados, C.A., Gaglia, J.L., Ryu, A., Waldner, H.,  
 Chernova, T., Manning, S., Greenfield, E.A., Coyle, A.J., Sobel, R.A.,  
 Freeman, G.J. and Kuchroo, V.K.  
 T11-specific cell surface protein Tim-3 regulates macrophage  
 activation and severity of an autoimmune disease  
 JOURNAL Nature 415 (6871), 536-541 (2002)  
 MEDLINE 2168188  
 PUBMED 11823861  
 REFERENCE  
 2 (bases 1 to 2725)  
 Monney, L., Sabados, C., Gaglia, J.L., Ryu, A., Waldner, H.,  
 Chernova, T., Greenfield, E.A., Sobel, R.A., Freeman, G.J. and  
 Kuchroo, V.K.  
 Direct Submision  
 Submitted (23-NOV-2001) Adult Oncology, Dana-Farber Cancer  
 Institute, 44 Binney St., Boston, MA 02115, USA  
 FEATURES  
 source  
 1..2725  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="AKR"  
 /db\_xref="taxon:10090"  
 /cell\_type="T cell"  
 67..312  
 /note="TIM-3"  
 /codon\_start=1  
 /product="T cell immunoglobulin mucin-3"  
 /protein\_id="AA165156.1"  
 /db\_xref="gi:18182531"  
 /translation="MPSGRTNLCVLLQLLLASRLDNGYVYEVKNAYLPCSYTLPT  
 SGLIVPMWKGKCPWQCTNLELRIBRNVTYQSSRYQLKGLDKADVSLIKVLT  
 LDDGTYCCRIQFPLNWDKLEIKDIKAQVPAQTAHSDSTLASRRLITERNGS  
 ETQTLVTLHNNGTKISTWADRIDSEITRTAIIHGVASAGLTLLAIIGVLILKMY  
 SCKKKKTSLSLITLAILPQGLNAGAVRISSENIYTIIEENVYEVENSNEYCYVN  
 SQQPS"  
 BASE COUNT 805 a 587 c 641 g 692 t  
 ORIGIN  
 Query Match 96.8%; Score 2624.6; DB 10; length 2725;  
 Best Local Similarity 99.0%; Pred. No. 0;







QY 2178 GGTGACGGGACATACGAAGGAGAGCTCTTGACCAATCTATATATAGTCAGCAAAATATT 2237  
DB 2203 GGATGACTGGGACATACGAAGGAGAGCTCTTGACCAATCTATATATAGTCAGCAAAATATT 2262  
QY 2228 CTTTGGGAGGACAGTCGTCCCAATTTGATTTCCAGCGGTGAGCTTCAGTTTATCTG 2297  
DB 2263 CTTTGGGAGGACAGTCGTCCCAATTTGATTTCCAGCGGTGAGCTTCAGTTTATCTG 2322  
QY 2298 GCTTACAGCTGCTGCTCCAGTGCCTTTGATCTGTGCTGGCTCCATCTATAACAGATCA 2357  
DB 2323 GCTTACAGCTGCTGCTCCAGTGCCTTTGATCTGTGCTGGCTCCATCTATAACAGATCA 2382  
QY 2358 AATTAATATGACCCCGAGTGAATAATTAAGTGAAGACAGAAAGTACTTTGTTCAAGAT 2417  
DB 2383 AATTAATATGACCCCGAGTGAATAATTAAGTGAAGACAGAAAGTACTTTGTTCAAGAT 2442  
QY 2418 TTTTTCGATTTGGGAGACATCTGTATCATCAGAGACATCTGTTAGTGAAGACCCAA 2477  
DB 2443 TTTTTCGATTTGGGAGACATCTGTATCATCAGAGACATCTGTTAGTGAAGACCCAA 2502  
QY 2478 ACCGTGGTACCGTTTTCATGATGATTAATTTGTTTGGTGGCTTCTAGCTAGCTG 2537  
DB 2503 ACCGTGGTACCGTTTTCATGATGATTAATTTGTTTGGTGGCTTCTAGCTAGCTG 2562  
QY 2538 TGAAGTCTGAGCTTTCTTAAAGTGGATGAGAGGAGACCATCTAACAATAATCCATTAG 2597  
DB 2563 TGAAGTCTGAGCTTTCTTAAAGTGGATGAGAGGAGACCATCTAACAATAATCCATTAG 2622  
QY 2598 AGATACAGCTCTCATGCAAGGAGAACTAATCTCAATGTTTAAAGTAAATAACT 2657  
DB 2623 AGATACAGCTCTCATGCAAGGAGAACTAATCTCAATGTTTAAAGTAAATAACT 2682  
QY 2658 GTACGGGCAAGTACTTTGAGCATTAATAAAAAAAAAAAAAA 2700  
DB 2683 GTACGGGCAAGTACTTTGAGCATTAATAAAAAAAAAAAAAA 2725

RESULT 6  
AL669948  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AL669948 223030 bp DNA linear ROD 05-APR-2002  
Mouse DNA sequence from clone RP23-248K2 on chromosome 11, complete  
sequence.  
AL669948  
AL669948.7 GI:19772854  
HTG.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.  
1  
Tracey A.  
Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humukey@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
On Mar 27, 2002 this sequence version replaced gi:1952137.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP  
database can be found at

QY 719 GTTGAAGCTTATACATGAGCCAACTGCTCCAGAGGGGTTGGCAATGACAGAGCAG 778  
DB 55170 GTTGAAGCTTATACATGAGCCAACTGCTCCAGAGGGGTTGGCAATGACAGAGCAG 55229  
QY 779 TCAGATTCGCTGAGAGAAATATCTACACCATGAGAGAAAGTATATGAGTGAAGA 838  
DB 55220 TCAGATTCGCTGAGAGAAATATCTACACCATGAGAGAAAGTATATGAGTGAAGA 55289  
QY 839 ATTCAATGAGTACTACTGCTACGTCACACGACGACGACCATCTGACCGCTCTGACT 898  
DB 55290 ATTCAATGAGTACTACTGCTACGTCACACGACGACGACCATCTGACCGCTCTGACT 55349  
QY 899 GCCACTTTAAAGCTGCGCTTCATTTCTGATTTGATTTCTCTTTTGTGAAAAATAT 958  
DB 55350 GCCACTTTAAAGCTGCGCTTCATTTCTGATTTGATTTCTCTTTTGTGAAAAATAT 55408  
QY 959 GTGATATGCTCACTTGGCAACCTCATTTGAGAGGTTCTGACACAGCCACTGAGAAAAAGTT 1018  
DB 55409 GTGATATGCTCACTTGGCAACCTCATTTGAGAGGTTCTGACACAGCCACTGAGAAAAAGTT 55468  
QY 1019 CGAGTTTCTGGGAGTAATTAATCTACCAAGGGGATTCGACTGTAACTCATGCTACATTGA 1078  
DB 55469 CGAGTTTCTGGGAGTAATTAATCTACCAAGGGGATTCGACTGTAACTCATGCTACATTGA 55528  
QY 1079 AATGCTCATTTTATCCCTGAGTTTCAGGAGTCCGATCCCACTCCAGACACTTCAATC 1138  
DB 55529 AATGCTCATTTTATCCCTGAGTTTCAGGAGTCCGATCCCACTCCAGACACTTCAATC 555  
QY 1139 ATGCGTTTGAAGCTCACTCGTCTTCTATCATTAAGGATGTAGTGTGCTTTG 1198  
DB 55589 ATGCGTTTGAAGCTCACTCGTCTTCTATCATTAAGGATGTAGTGTGCTTTG 55648  
QY 1199 AGACATAGAGGTTTGTGTATATCCGCAAGCTCTGAAACAGTGAAGGGGATAAAGGC 1258  
DB 55649 AGACATAGAGGTTTGTGTATATCCGCAAGCTCTGAAACAGTGAAGGGGATAAAGGC 55708  
QY 1259 TAAATAGAGAGGTCGCTCTTTGTGTATGTGAAAAATCTTAAAGAGTGTGCTTT 1318  
DB 55709 TAAATAGAGAGGTCGCTCTTTGTGTATGTGAAAAATCTTAAAGAGTGTGCTTT 55767  
QY 1319 TCTAGAGATTTCTGACCTTGAAGAGTAAAGAAAAAGCAGAGGGGATGTGCTTACACGA 1378  
DB 55768 TCTAGAGATTTCTGACCTTGAAGAGTAAAGAAAAAGCAGAGGGGATGTGCTTACACGA 55827  
QY 1379 TATAACTTGGAGACCTTGAAGAGGAGGTGATTAAGTGAAGTGAAGGAGGCTTATGCTG 1438  
DB 55828 TATAACTTGGAGACCTTGAAGAGGAGGTGATTAAGTGAAGTGAAGGAGGCTTATGCTG 55887  
QY 1439 GTAAGACTGTCTCMCAATCCAAAGCAAAATTAACATATGAGACAGCAGAGGCTGAGGA 1498  
DB 55888 GTAAGACTGTCTCMCAATCCAAAGCAAAATTAACATATGAGACAGCAGAGGCTGAGGA 55947  
QY 1499 TGAAGCTCGACAGTGAAGTGCATTTGTATACAGACAGGAATCTATATTTGATCGTAG 1558

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep/RP23-248K2 is  
from the RPCI-23 Mouse PAC Library  
constructed by the group of Pieter de Jong.  
For further details see http://www.chori.org/bacpac/home.htm  
VECTOR: pBACs.6.  
Location/Qualifiers  
1. 223030  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/clone="RP23-248K2"  
/clone\_11b="RPCI-23"

BASE COUNT 63014 a 47122 c 47217 g 65677 t  
ORIGIN  
Query Match 69.6%; Score 1886.6; DB 10; Length 223030;  
Best Local Similarity 97.7%; Pred. No. 0;  
Matches 1941; Conservative 6; Mismatches 30; Indels 10; Gaps 3;

QY 719 GTTGAAGCTTATACATGAGCCAACTGCTCCAGAGGGGTTGGCAATGACAGAGCAG 778  
DB 55170 GTTGAAGCTTATACATGAGCCAACTGCTCCAGAGGGGTTGGCAATGACAGAGCAG 55229  
QY 779 TCAGATTCGCTGAGAGAAATATCTACACCATGAGAGAAAGTATATGAGTGAAGA 838  
DB 55220 TCAGATTCGCTGAGAGAAATATCTACACCATGAGAGAAAGTATATGAGTGAAGA 55289  
QY 839 ATTCAATGAGTACTACTGCTACGTCACACGACGACGACCATCTGACCGCTCTGACT 898  
DB 55290 ATTCAATGAGTACTACTGCTACGTCACACGACGACGACCATCTGACCGCTCTGACT 55349  
QY 899 GCCACTTTAAAGCTGCGCTTCATTTCTGATTTGATTTCTCTTTTGTGAAAAATAT 958  
DB 55350 GCCACTTTAAAGCTGCGCTTCATTTCTGATTTGATTTCTCTTTTGTGAAAAATAT 55408  
QY 959 GTGATATGCTCACTTGGCAACCTCATTTGAGAGGTTCTGACACAGCCACTGAGAAAAAGTT 1018  
DB 55409 GTGATATGCTCACTTGGCAACCTCATTTGAGAGGTTCTGACACAGCCACTGAGAAAAAGTT 55468  
QY 1019 CGAGTTTCTGGGAGTAATTAATCTACCAAGGGGATTCGACTGTAACTCATGCTACATTGA 1078  
DB 55469 CGAGTTTCTGGGAGTAATTAATCTACCAAGGGGATTCGACTGTAACTCATGCTACATTGA 55528  
QY 1079 AATGCTCATTTTATCCCTGAGTTTCAGGAGTCCGATCCCACTCCAGACACTTCAATC 1138  
DB 55529 AATGCTCATTTTATCCCTGAGTTTCAGGAGTCCGATCCCACTCCAGACACTTCAATC 555  
QY 1139 ATGCGTTTGAAGCTCACTCGTCTTCTATCATTAAGGATGTAGTGTGCTTTG 1198  
DB 55589 ATGCGTTTGAAGCTCACTCGTCTTCTATCATTAAGGATGTAGTGTGCTTTG 55648  
QY 1199 AGACATAGAGGTTTGTGTATATCCGCAAGCTCTGAAACAGTGAAGGGGATAAAGGC 1258  
DB 55649 AGACATAGAGGTTTGTGTATATCCGCAAGCTCTGAAACAGTGAAGGGGATAAAGGC 55708  
QY 1259 TAAATAGAGAGGTCGCTCTTTGTGTATGTGAAAAATCTTAAAGAGTGTGCTTT 1318  
DB 55709 TAAATAGAGAGGTCGCTCTTTGTGTATGTGAAAAATCTTAAAGAGTGTGCTTT 55767  
QY 1319 TCTAGAGATTTCTGACCTTGAAGAGTAAAGAAAAAGCAGAGGGGATGTGCTTACACGA 1378  
DB 55768 TCTAGAGATTTCTGACCTTGAAGAGTAAAGAAAAAGCAGAGGGGATGTGCTTACACGA 55827  
QY 1379 TATAACTTGGAGACCTTGAAGAGGAGGTGATTAAGTGAAGTGAAGGAGGCTTATGCTG 1438  
DB 55828 TATAACTTGGAGACCTTGAAGAGGAGGTGATTAAGTGAAGTGAAGGAGGCTTATGCTG 55887  
QY 1439 GTAAGACTGTCTCMCAATCCAAAGCAAAATTAACATATGAGACAGCAGAGGCTGAGGA 1498  
DB 55888 GTAAGACTGTCTCMCAATCCAAAGCAAAATTAACATATGAGACAGCAGAGGCTGAGGA 55947  
QY 1499 TGAAGCTCGACAGTGAAGTGCATTTGTATACAGACAGGAATCTATATTTGATCGTAG 1558

Db 55948 TGAAGTCCGACGATGAGTGAATTTGTAACGACGAGAAATCTATATTTGATGTAG 56007  
 QY 1559 ACCCCACATGAAAAAGCTAGGCTGTAGAGCATGCTTTAGACTCAAGAGATGAGAG 1618  
 Db 56008 ACCCCACATGAAAAAGCTAGGCTGTAGAGCATGCTTTAGACTCAAGAGATGAGAG 56067  
 QY 1619 TAAAGGACACAGATCCCGGGGGCTTGGGTGACATGCTTACCTAGGTGCTGATTG 1678  
 Db 56068 TAAAGGACACAGATCCCGGGGGCTTGGGTGACATGCTTACCTAGGTGCTGATTG 56127  
 QY 1679 CAAGTCCACAAAGAGTCCCTGTCTCAAGATGATGATGATGATGATGATGATGATGAT 1738  
 Db 56128 CAAGTCCACAAAGAGTCCCTGTCTCAAGATGATGATGATGATGATGATGATGATGAT 56187  
 QY 1739 GGGGGTGTCTCTCTCTCTCAGAAAGACATGACATGACATGACATGACATGACATGACAT 1798  
 Db 56188 GGGGGTGTCTCTCTCTCTCAGAAAGACATGACATGACATGACATGACATGACATGACAT 56247  
 QY 1799 CAC 1850  
 Db 56248 CAC 56307  
 QY 1851 TGCTGTCT 1910  
 Db 56308 TGCTGTCT 56367  
 QY 1911 ATGAGTGGAGACATGAC 1970  
 Db 56368 ATGAGTGGAGACATGAC 56427  
 QY 1971 AGAAGGTGAGAGATTCAGAGACATGCTGCTCCCAACCTGCGACATCTCTTCAACAAGA 2030  
 Db 56428 AGAAGGTGAGAGATTCAGAGACATGCTGCTCCCAACCTGCGACATCTCTTCAACAAGA 56487  
 QY 2031 AAAAGGTATCCCGACAGACACCTGCTGTCTGTAAGAAACCTGCGAAAGGCAAACTT 2090  
 Db 56488 AAAAGGTATCCCGACAGACACCTGCTGTCTGTAAGAAACCTGCGAAAGGCAAACTT 56547  
 QY 2091 TGAAGTGTGTGCTCAGAGGGAATGACTCAGACAACTTCTCTCTCTCTCTCTCTCTCTCT 2150  
 Db 56548 TGAAGTGTGTGCTCAGAGGGAATGACTCAGACAACTTCTCTCTCTCTCTCTCTCTCTCT 56607  
 QY 2151 GGAAGCTTTTTCAGAGAAAGAAACACCGGTGATGCGACATGAGAAAGGCAAGCTTTTGC 2210  
 Db 56608 GGAAGCTTTTTCAGAGAAAGAAACACCGGTGATGCGACATGAGAAAGGCAAGCTTTTGC 56667  
 QY 2211 AGCAATCTATATGATGAGCAAAATATTTTGGGAGAGACATGCTCACCACAAATGATTTG 2270  
 Db 56668 AGCAATCTATATGATGAGCAAAATATTTTGGGAGAGACATGCTCACCACAAATGATTTG 56727  
 QY 2271 CAAGCGGTGAGACCTGAGTTTCATCTGCTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCT 2330  
 Db 56728 CAAGCGGTGAGACCTGAGTTTCATCTGCTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCT 56787  
 QY 2331 TGTGTGCTCCCATCTATTAACAGATCAATTAATGACCCCGAGTGAATATTTAAGT 2390  
 Db 56788 TGTGTGCTCCCATCTATTAACAGATCAATTAATGACCCCGAGTGAATATTTAAGT 56847  
 QY 2391 ACCAGAAAGTATGCTTTTCAAGATTTTTTTCATTTGGGAGAGCAATGCTGATCATGAC 2450  
 Db 56848 ACCAGAAAGTATGCTTTTCAAGATTTTTTTCATTTGGGAGAGCAATGCTGATCATGAC 56907  
 QY 2451 AGGACATCTGTATGATGAGGACACCAAAACCTGTGTGATACCGTTTTCATGTATGAATTT 2510  
 Db 56908 AGGACATCTGTATGATGAGGACACCAAAACCTGTGTGATACCGTTTTCATGTATGAATTT 56967  
 QY 2511 GTTGTATGATGCTCTCTGAGTGTGAGAGCTGCTGCTTCTTATAGGTGATGAGAA 2570  
 Db 56968 GTTGTATGATGCTCTCTGAGTGTGAGAGCTGCTGCTTCTTATAGGTGATGAGAA 57027  
 QY 2571 GGGAGACATCTAACAAATCCATTTAGAGATTAACAGCTCTCATGCAAGAGGAAACTTAA 2630

Db 57028 GGGAGACCATCTAACAAATCCATTTAGAGATTAACAGCTCTCATGCAAGAGGAAACTTAA 57087  
 QY 2631 TCTCAATGTTTTTAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2690  
 Db 57088 TCTCAATGTTTTTAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 57147  
 QY 2691 AAAAAA 2697  
 Db 57148 AGATCA 57154  
 RESULT 7  
 AC119505/C  
 LOCUS  
 DEFINITION  
 Rattus norvegicus clone CH230-344M17, WORKING DRAFT SEQUENCE, 3  
 UNORDERED PICES.  
 AC119505  
 AC119505.5 GI:25138135  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULIDOP.  
 KEYWORDS  
 Rattus norvegicus (Norway rat)  
 SOURCE  
 ORGANISM  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 177087)  
 Muzny,D,Marie, Metzker,M, Lee, A, Abramson, S, Adams, C, Alder, J, Allen, C, Allen, H, Alsdorfs, S, Amth, A, Angilano, D, Anyalebechi, V, Ayogei, A, Ayodeji, M, Baca, E, Baden, H, Baldwin, D, Bandaranaike, D, Barber, M, Barnstead, M, Benahmed, F, Biswal, K, Blair, J, Blankenburg, K, Blyth, P, Brown, M, Bryant, N, Buhay, C, Burch, P, Burrell, K, Calderon, E, Cardenas, V, Carter, K, Cavazos, I, Ceasar, H, Center, A, Chacko, J, Chavez, D, Chen, G, Chen, R, Chen, Y, Chen, Z, Chu, J, Cleveland, C, Cockrell, R, Cox, C, Coyle, M, Cree, A, D'Souza, L, Davila, M, L, Davis, C, Davy-Carroll, L, De Anda, C, Dederich, D, Delgado, O, Denison, S, Deramo, C, Ding, Y, Dinh, H, Divya, K, Draper, H, Dugan-Rocha, S, Dunn, A, Durbin, K, Duval, B, Eaves, K, Egan, A, Becotto, M, Eugene, C, Evans, C, A, Falls, T, Fan, G, Fernandez, S, Finley, M, Flagg, N, Forbes, L, Foster, P, Fraser, C, M, Gabisi, A, Ganta, R, Garcia, A, Garner, T, Garza, M, Gebregiorgis, E, Geer, K, Gill, R, Grady, M, Guerra, W, Guevara, W, Gunaratne, P, Haaland, W, Hamill, C, Hamilton, C, Hamilton, K, Harvey, Y, Havlik, P, Hawes, A, Henderson, N, Hernandez, J, Hernandez, R, Hines, S, Hladun, S, L, Hodgson, A, Hughes, M, Hollins, B, Howells, S, Huylk, S, Hune, J, Idlebird, D, Jackson, A, Jackson, L, Jacob, L, Jiang, H, Johnson, B, Johnson, R, Jolivet, A, Karpathy, S, Kelly, S, Kelly, S, Khan, Z, King, L, Kovar, C, Kowis, C, Kraft, C, L, Lebow, H, Levan, J, Lewis, L, Li, Z, Liu, J, Liu, J, Liu, W, Liu, Y, London, P, Longacre, S, Lopez, J, Lorenshewe, L, Louised, H, Lozano, R, J, Lu, X, Ma, J, Maheshwari, M, Mahindaratne, M, Mahmood, M, Malloy, K, Mangum, A, Mangum, B, Mapua, P, Martin, K, Martin, R, Martinez, E, Mawhinney, S, McLeod, M, P, McNeill, T, Z, Meenen, E, Milosavljevic, A, Miner, G, Minja, E, Montemayor, J, Morgan, M, Morris, K, Morris, S, Mundiase, M, Murphy, M, Nat, L, Nankervill, C, Neal, D, Newton, N, Nguyen, N, Norris, S, Parks, K, Nwakoelimen, O, Okumou, G, Olarunwasagoun, A, Pal, S, Parks, K, Pasternak, S, Paul, H, Perez, A, Perez, L, Pfamkocho, C, Plopper, F, Poindexter, A, Popovic, D, Primus, E, Pu, L, L, Puazo, M, Quiroz, J, Rachlin, E, Reeves, K, Regier, M, Reigh, R, Reilly, B, Reilly, M, Ren, Y, Reuter, M, Richards, S, Riggs, F, Rivers, C, Rodkey, T, Rojas, A, Rose, M, Rose, R, Ruiz, S, J, Sanders, W, Savery, G, Scherer, S, Scott, G, Shatman, S, Shen, H, Shetty, J, Shvartsbeyn, A, Sisson, I, Sitter, C, D, Smjs, D, Sneed, A, Sodergren, E, Song, X, Z, Sorelle, R, Sosa, J, Steidle, M, Strong, R, Sutton, A, Svatek, A, Taber, P, Taylor, C, Taylor, T, Thomas, N, Thomas, S, Tinney, A, Trejos, Z, Usmani, K, Valas, R, Vera, V, Villaseana, D, Waldron, L, Walker, B, Wang, J, Wang, Q, Wang, S, Warren, J, Warren, R, Wei, X, White, F, Williams, G, Willson, R, Wleczky, R, Woodson, H, Woley, K, Wright, D, Wright, R, Wu, J, Yakub, S, Yen, J, Yoon, L, Yoon, V, Yu, F, Zhang, J, Zhou, J, Zhou, X, Zhao, S, Dunn, D, von Niederhausen, A, Weiss, R, Smith, D, R, Holt, R, A, Smith, H, O,







Db 191387 GTGCATGCTTGTGACCTCCAGAGATAGGAGGTAAAGGCAAGACATGCCCTGGGGCATG 191328  
 QY 1647 CGTGACGTAGCTTACCTTAGCTGCTGAGTTCACAGAGATCCCTGTCTCAGAG 1706  
 Db 191327 ATTGAGCCACCTTACCTTAGAGTGTGAGTTCAGAGATCCCTGTCTCAGAT 191268  
 QY 1707 TAAGATGAGTGTGATCTGAGGCGCATGTCCATGGGGGTGTCTCTCTCTCAGAG--- 1763  
 Db 191267 TAAGATGAGTGTGATCTGAGGCGCATGTCCATGGAGATCTCTCTCTCTCTCAGAG 191208  
 QY 1764 -----AGACATGACCATGTGCCCTGACACACACACACACACACACACAC 1809  
 Db 191207 TGGCATGTACCTTACACACACACATGCGCATGCGACCGCAATATGCGCACATG 191148  
 QY 1810 ACACACACACACACACACATGAMA-----TGAG 1840  
 Db 191147 GGTACACCCCAAAAGATTAAAGAAAGATCAGAACGGGCTAAATGATAGTGA 191088  
 QY 1841 GTTCTCTGTGCTGCTGACCTCTCTATTAACATGTATCTCTACAGGACTCTCTGCGCT 1900  
 Db 191087 GTTCTCTGTGCTGCTGACCTCTCTATTAATGTTTATCTCTATATAGCTCTCTCTGCGCT 191028  
 QY 1901 CTGTTAAGACA-----TGAGTGAGACATGCGCAGACAGTCCAGTAATTATTCAGAC 1955  
 Db 191027 CTGTTAAGACTGCTCTGAAACGGAGACACTGAATGACACTGTATTAATTCAGAC 190968  
 QY 1956 TCGAAGGCTGAGCGAGAACCGTGTGAGATTCAGAGACTGTGCCACACTGCGCAC 2015  
 Db 190967 CCGAAGGTTGAGCGAGACATGAGAGAGTCCAGAGCATGTGTCTTACACTGCGCAC 190908  
 QY 2016 TCTTTCTTACACAGAAAAGGTTTACCGGACAGCGCTGTCTGTAAAGGAAAC--C 2073  
 Db 190907 TC-TCTTACACAGAAAAGGTTTACCGGACATGAGTGTGTCTGTGAGGAAAGAACTTT 190849  
 QY 2074 CTGCGAAGCGCAACTTTGACTGTGTGTGCTCAGAGGGAACCTACTCAGACACTTCTC 2133  
 Db 190848 TTAAGAACTGACCTGCTGCTGTGTGTCTCAGAGGGAACCTACTCAGACACTTCTC 190789  
 QY 2134 CATTCTGTGAGGAACTGTGAGC--TGTTTCTGACAGAAACACCGGTGACTGCGGACAT 2191  
 Db 190788 CATTCTGTGAGGAACTGTGAGCCTGTGTCTGACAGAAAGAACCGGTGACTGCGGACAC 190729  
 QY 2192 ACAGAGCGAGAGCTTGTGACGACATCTATTAAGTACGACAAATATTC--TTTGGAGACA 2250  
 Db 190728 AGAGAGCGAGAGCTTGTGACGACATCTATTAAGTACGACAAAGATCTTTTGGTAGACA 190669  
 QY 2251 GTCTGTACCAAA--TTGATTTCCAGCGCGTGAACCTCAGTTTCATCTGCGTTACAGCTGC 2309  
 Db 190668 GTGTGTACCAAACTTAAATTTCCAGCGGTGTGTGAGCTCAGTTTCATCTGCGTTACAGCTGC 190609  
 QY 2310 CTGCCAGTGCCTTGTATCTGTCTGTGCTCCATCTATAACAGATCAATTAATATAGC 2369  
 Db 190608 TTGGCCAGTG--CCTGTGTGTGTGTGTCTCAATAATAACAGATC--AATTAATATAGC 190551  
 QY 2370 CCCGAGTGAATAATTAAGTACAGAGAGTGTGTTGTTTCAAGATTTTTTGTGATG 2429  
 Db 190550 CCGAGTGAATAATTAAGTACAGAGAGTGTGTTGTTTCAAGATTTTTTGTGATG 190492  
 QY 2430 GGGAGCACTGTGTACATCAGAGACATCTGTATGAGAGACCAAAACCTG--TGTA 2487  
 Db 190491 GGGTGTACTGTGTATGTCAAGAGACGTCTGTATGAGAGCACTGAAACCTCAGTTT 190432  
 QY 2488 CCGTTTTTTCATGTATGAAATTTTGTGTGTTAGTGTCTTCTAGCTAGCTGTGAAGTCTCT 2547  
 Db 190431 GGAATCTTTTCTGTATGAAATTTTGTGTGTTAGTGTCTCATTAGCTGTGAAGTCTCT 190372  
 QY 2548 GCTTTCTTAGTGTATGAGAGGAGACATCTAACAATAATCAATTAGAGATTAACAG 2607  
 Db 190371 GGAATCTTTTAGTGTATGAGAGGAGACATCTCAGAAATCAAT--AGGTAACATC 190315  
 QY 2608 TCTCATGACAGAGAGAGAACTAAATCTCAATGTTTTTAAAGTAATA--AACTGTACTG 2663  
 Db 190314 TCTCATGACAGAGAGAGAACTAAATCTCAAGTGTGTTTTAAAGTAATAACCAACTGTACT 190255

QY 2664 GCMAAGTACTTTGACATTAATAAAAAAAAAAAAAA 2697  
 Db 190254 GTGAAGTACTTTGACATTAATAATAACAGATCA 190221  
 RESULT 9  
 AF399831  
 LOCUS AF399831 862 bp mRNA linear ROD 08-APR-2002  
 DEFINITION Mus musculus strain DBA/2 TIM3 mRNA, partial cds.  
 ACCESSION AF399831  
 VERSION AF399831.1 GI:17148680  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 REFERENCE  
 AUTHORS McIntire,J.J., Umetsu,S.E., Akbari,O., Potter,M., Barsh,G.S.,  
 1 (bases 1 to 862)  
 TITLE Identification of Tgpr (an airway hyperreactivity regulatory locus) and the linked Tim gene family  
 JOURNAL Nat. Immunol. 2 (12), 1109-1116 (2001)  
 MEDLINE 21582130  
 PUBMED 11725301  
 REFERENCE  
 AUTHORS McIntire,J.J., Umetsu,S.E., Akbari,O., Potter,M., Barsh,G.S.,  
 2 (bases 1 to 862)  
 TITLE Direct Submission  
 JOURNAL Submitted (12-JUN-2001) Pediatrics, Stanford University, 300  
 Pasteur Drive, Grant S-303, Stanford, CA 94305, USA  
 FEATURES  
 source location/Qualifiers  
 1..862  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="DBA/2"  
 /db\_xref="taxon:10090"  
 /chromosome="11"  
 /map="between D1Mit271 and D1Mit22"  
 /tissue\_type="spleen"  
 17..862  
 /note="contains T-cell immunoglobulin domain and mucin domain"  
 /codon\_start=1  
 /product="TIM3"  
 /protein\_id="AAL35776.1"  
 /db\_xref="GI:17148681"  
 /translation="MPSGLTNCVLLQLLARSLENAYVEGKNAYLPCSYLST  
 PGALVPMCKGKGFCEWSCQTNELRTDERNTVYKRSRYLKGDLKGVSLIKNVT  
 LDDHGTVCRIOPGILNDKLELKDIAKATVPAQTAHDSPTASPTLTRENGE  
 ETQTLVTHNNNGTKISTWADETKDSGETIRTAHITGVGSAGLTALITGLIKRY  
 SCKRRKLSLSLITLANLPGLAAGAVRIKRSSENIYITSENYEVENSNKYCVN  
 SOQPS"  
 BASE COUNT 254 a 197 c 198 g 213 t  
 ORIGIN  
 Query Match 31.2%; Score 844.4; DB 10; Length 862;  
 Best Local Similarity 98.7%; Pred. No. 2,1e-233;  
 Matches 851; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 QY 24 CCCCTCCCAAGTACATGTTTTCAGGTCTTACGCTCACTGCTGCTGCTGCA 83  
 Db 1 CCCCTCCCAAGTACATGTTTTCAGGTCTTACGCTCACTGCTGCTGCTGCA 60  
 QY 84 ACTACTACTTGGAGAGTCAATGGAAGATGTTATTAAGGTGAGGTGTGTAATGCTTA 143  
 Db 61 ACTACTACTTGGAGAGTCAATGGAAGATGTTATGTTTGAAGTGAAGAAAGCTTA 120  
 QY 144 TCTGCCCTGCACTTACACTTACATCTGAGACACTTGTGCTTATGCTGGGCGCA 203  
 Db 121 TCTGCCCTGCACTTACACTTACATCTGAGGCACTTGTGCTTATGCTGGGCGCA 180  
 QY 204 GGGATTCTGTCTTGTGTCACTGTACCAATGAGTTGCTCAGAACTGATGAAGAAATGT 263

```

Db 181 GGGATTTCCTCTGGTCAAGTGTACCAACGAGTTGCTGAGAACTGATGAAGAAATGT 240
Qy 264 GACATATGAAATGACAGATACACAGTAAAGGCGATCTCAACAAAGAGATGTGTC 323
Db 241 GACATATGAAATGACAGATACACAGTAAAGGCGATCTCAACAAAGAGATGTGTC 300
Qy 324 TCTGATCATAAAGATGTGACTGTGATGACCATGGGACTTACTGTGAGATACAGTT 383
Db 301 TCTGATCATAAAGATGTGACTGTGATGACCATGGGACTTACTGTGAGATACAGTT 360
Qy 384 CCTGTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443
Db 361 CCTGTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Qy 444 CACTCCACTGAGTACGCGCATGGGACTTACTGACTTCTCCAAAGACCTTAACAC 503
Db 421 CACTCCACTGAGTACGCGCATGGGACTTACTGACTTCTCCAAAGACCTTAACAC 480
Qy 504 GGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 563
Db 481 GGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Qy 564 AATTTCACATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 623
Db 541 AATTTCACATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Qy 624 CATTGAGATGGAGCTGCTGCTGGGTTGACCTGCGACTTATCATTTGCTTATCTCT 683
Db 601 CATTGAGATGGAGCTGCTGCTGGGTTGACCTGCGACTTATCATTTGCTTATCTCT 660
Qy 684 TAAATGATATTCCTGTAAGAAAAAGATTAATGAGTTGAGCTTATTAACCTGGCAA 743
Db 661 TAAATGATATTCCTGTAAGAAAAAGATTAATGAGTTGAGCTTATTAACCTGGCAA 720
Qy 744 CTGCTCTCAGAGAGGTTGGCAATGACAGAGCACTGAGATTGCTCTGAGAAAAATAT 803
Db 721 CTGCTCTCAGAGAGGTTGGCAATGACAGAGCACTGAGATTGCTCTGAGAAAAATAT 780
Qy 804 CTACACCATCGAGAGAGATGATATGATGAGTGAATCAATGATGATGATGATGATGAT 863
Db 781 CTACACCATCGAGAGAGATGATATGATGAGTGAATCAATGATGATGATGATGATGAT 840
Qy 864 CAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 885
Db 841 CAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 862

```

```

RESULT 10
ARI43567 843 bp DNA linear PAT 08-AUG-2001
LOCUS ARI43567
DEFINITION Sequence 22 from patent US 6204371.
ACCESSION ARI43567
VERSION ARI43567.1 GI:15104853
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 843)
AUTHORS Levinson, D.Adam.
TITLE Compositions and methods for the treatment and diagnosis of immune disorders
JOURNAL Patent: US 6204371-A 22-20-MAR-2001;
FEATURES
SOURCE Location/Qualifiers
BASE COUNT 253 a 185 c 196 g 209 t
ORIGIN

```

Query Match 31.1%; Score 843; DB 6; Length 843;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-233;  
 Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 40 ATGTTTCAGGCTTACCTCAACGTGTCTGCTGCTGCTGCAACTACTACTTCAAG 99
Db 1 ATGTTTCAGGCTTACCTCAACGTGTCTGCTGCTGCTGCAACTACTACTTCAAG 60
Qy 100 TCATTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 159
Db 61 TCATTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Qy 160 ACTTCACTATATCTGGGACCTGTGTGCTATGTGCTGGGCAAGGATTTCTGCTTGG 219
Db 121 ACTTCACTATATCTGGGACCTGTGTGCTATGTGCTGGGCAAGGATTTCTGCTTGG 180
Qy 220 TCACAGTGTACCAATGAGTTGCTCAGAACTGATGAAAGAAATGTGACATATCAGAAATCC 279
Db 181 TCACAGTGTACCAATGAGTTGCTCAGAACTGATGAAAGAAATGTGACATATCAGAAATCC 240
Qy 280 AGCAGATACGAGCTAAAGGGGATCTCAACAAAGAGATGTGTCTGTGATTAAGAAAT 339
Db 241 AGCAGATACGAGCTAAAGGGGATCTCAACAAAGAGATGTGTCTGTGATTAAGAAAT 300
Qy 340 GTGACTCTGATGACATGAGGACCTTACTGCTGCAAGATACAGTTCCCTGCTTATGAT 399
Db 301 GTGACTCTGATGACATGAGGACCTTACTGCTGCAAGATACAGTTCCCTGCTTATGAT 360
Qy 400 GATAAAAAATTAGAACTGAAATTAGACATCAAGCAGCAAGTCACTCCAGCTCAGACT 459
Db 361 GATAAAAAATTAGAACTGAAATTAGACATCAAGCAGCAAGTCACTCCAGCTCAGACT 420
Qy 460 GCCCATGGGAGCTTACTACAGCTTCTCAAGAAACCTTAACACGAGAGAAATGTTCA 519
Db 421 GCCCATGGGAGCTTACTACAGCTTCTCAAGAAACCTTAACACGAGAGAAATGTTCA 480
Qy 520 GAGACACAGACATGAGTGAACCTCCATTAATTAAGATGAACAAATTTCCACATGGGCT 579
Db 481 GAGACACAGACATGAGTGAACCTCCATTAATTAAGATGAACAAATTTCCACATGGGCT 540
Qy 580 GATGAAATTAAGAACTGAGAGAAAGATGAGAACTGATATCCACATGAGTGAAGTGC 639
Db 541 GATGAAATTAAGAACTGAGAGAAAGATGAGAACTGATATCCACATGAGTGAAGTGC 600
Qy 640 TCTGCTGGGTGACCTTGGCACTTATCATTTGCTTAAATCTTAAATGATTTCTGT 699
Db 601 TCTGCTGGGTGACCTTGGCACTTATCATTTGCTTAAATCTTAAATGATTTCTGT 660
Qy 700 AAGAAAAAAGATGATGAGTTGAGCTTATTAACATGAGCAACTGCTCCAGAGAGG 759
Db 661 AAGAAAAAAGATGATGAGTTGAGCTTATTAACATGAGCAACTGCTCCAGAGAGG 720
Qy 760 TTGGCAATGAGAGAGCAGTACAGATTGCTCTGAGGAAATATCTACACCATGAGAG 819
Db 721 TTGGCAATGAGAGAGCAGTACAGATTGCTCTGAGGAAATATCTACACCATGAGAG 780
Qy 820 AACGTATATGAGTGAAGTGAATTCAAATGAGTACTTACTGCTCAACAGCAGCAGCA 879
Db 781 AACGTATATGAGTGAAGTGAATTCAAATGAGTACTTACTGCTCAACAGCAGCAGCA 840
Qy 880 TCC 882
Db 841 TCC 843

```

```

RESULT 11
ARI68936 843 bp DNA linear PAT 17-DEC-2001
LOCUS ARI68936
DEFINITION Sequence 22 from patent US 6288218.
ACCESSION ARI68936
VERSION ARI68936.1 GI:17905116
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 843)
AUTHORS Levinson, D.Adam.

```



TITLE	Compositions and methods for the treatment and diagnosis of immune disorders
JOURNAL	Patent: US 6288218-A 22 11-SEP-2001;
FEATURES	Location/Qualifiers
source	1. 843
BASE COUNT	/organism="unknown"
ORIGIN	253 a 185 c 196 g 209 t
Query Match	31.1%; Score 843; DB 6; Length 843;
Best Local Similarity	100.0%; Pred. No. 5.3e-233;
Matches	843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	40 ATGTTTTCAGGCTTACCCCTCAACCTGTGTCCTGCTGCTGCACTACTACTTCCAGG 99
DB	1 ATGTTTTCAGGCTTACCCCTCAACCTGTGTCCTGCTGCTGCACTACTACTTCCAGG 60
QY	100 TCATTGGAAGATGATTATPAAGGTTGAGTTGGTATAAAATGCCATATCTGCCCTGCAGTTAC 159
DB	61 TCATTGGAAGATGATTATPAAGGTTGAGTTGGTATAAAATGCCATATCTGCCCTGCAGTTAC 120
QY	160 ACTTACTCTCACTCTGGGCACTTTGTGCTATGTCTGTGGGGCAAGAGATTTGTCTCTTGG 219
DB	121 ACTTACTCTCACTCTGGGCACTTTGTGCTATGTCTGTGGGGCAAGAGATTTGTCTCTTGG 180
QY	220 TCACAGTGTACCAATGAGTGTGCTCGAAGCTGATGTAAGAAAGAAATGTGACATATCAGAAATCC 279
DB	181 TCACAGTGTACCAATGAGTGTGCTCGAAGCTGATGTAAGAAAGAAATGTGACATATCAGAAATCC 240
QY	280 AGCAGATACCCAGCTTAAAGGGCGATCTCAACAAAGAGATGTGTCTTGATCATAAAGAT 339
DB	241 AGCAGATACCCAGCTTAAAGGGCGATCTCAACAAAGAGATGTGTCTTGATCATAAAGAT 300
QY	340 GTGATCTCTGATGACCAATGGGACCTATCTGCTGCAAGATATCACTTCCCTGTCTTATGAT 399
DB	301 GTGATCTCTGATGACCAATGGGACCTATCTGCTGCAAGATATCACTTCCCTGTCTTATGAT 360
QY	400 GATAAAATTTGAAGTGAATTTAGATCAATCAAGAGAGCCAGAGTCACTCCAGCTCAGACT 459
DB	361 GATAAAATTTGAAGTGAATTTAGATCAATCAAGAGAGCCAGAGTCACTCCAGCTCAGACT 420
QY	460 GCCCATGGGGACTCTACTACAGCTTCTCCAGAACCTTAAACGAGGAGAAATGTTTCA 519
DB	421 GCCCATGGGGACTCTACTACAGCTTCTCCAGAACCTTAAACGAGGAGAAATGTTTCA 480
QY	520 GAGACACAGACCTGTGTGACCTCCATATAATTAATGGAACAAAAATTTCCATAGGGCT 579
DB	481 GAGACACAGACCTGTGTGACCTCCATATAATTAATGGAACAAAAATTTCCATAGGGCT 540
QY	580 GATGAAATTTAAGGACTGTGAGAAACGATCAGAACTGTATCCACATTGAGTGGAGATC 639
DB	541 GATGAAATTTAAGGACTGTGAGAAACGATCAGAACTGTATCCACATTGAGTGGAGATC 600
QY	640 TCTGCTGGGTTGACCTTGGCACTTATCATTTGTGTCTTATATCTTAAATGGTATTCCTGT 699
DB	601 TCTGCTGGGTTGACCTTGGCACTTATCATTTGTGTCTTATATCTTAAATGGTATTCCTGT 660
QY	700 AAGAAAAAAGATTTATCGAGTTTGAGCCTTATTAATCACTGGGCCAATCTGCTCCAGAGGG 759
DB	661 AAGAAAAAAGATTTATCGAGTTTGAGCCTTATTAATCACTGGGCCAATCTGCTCCAGAGGG 720
QY	760 TTGGCAAAATGCAGAGCAGTCAAGATTGCTCTGAGAAATATATCTACATTCGAGGAG 819
DB	721 TTGGCAAAATGCAGAGCAGTCAAGATTGCTCTGAGAAATATATCTACATTCGAGGAG 780
QY	820 AACGTATATGAAGTGAAGATTTCAATATGATATCTCTGCTACGTCAACAGCCAGAGCCCA 879
DB	781 AACGTATATGAAGTGAAGATTTCAATATGATATCTCTGCTACGTCAACAGCCAGAGCCCA 840
QY	880 TCC 882
DB	841 TCC 843

[illegible]



QY 760 TTGGCAATGCGAGAGCAGTGGATTGCTCTGAGAGAAATATCTACACCATCGAGAG 819  
DB 721 TTGGCAATGCGAGAGCAGTGGATTGCTCTGAGAGAAATATCTACACCATCGAGAG 780  
QY 820 AACGATATGAGAGTGGAGAAATTTCAATGAGTACTGCTTACGTCACAGCCGAGAGCA 879  
DB 781 AACGATATGAGAGTGGAGAAATTTCAATGAGTACTGCTTACGTCACAGCCGAGAGCA 840  
QY 880 TCC 882  
DB 841 TCC 843

RESULT 13  
BD157433 1772 bp DNA linear PAT 17-JAN-2003  
LOCUS BD157433  
DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
ACCESSION BD157433  
VERSION BD157433.1 GI:27863191  
KEYWORDS JP 2002191363-A/12276.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1772)  
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
Primer for synthesizing full-length cDNA and use thereof  
Patent: JP 2002191363-A 12276 09-Jul-2002;  
HELIX RESEARCH INSTITUTE  
OS Homo sapiens (human)  
PN JP 2002191363-A/12276  
PD 09-Jul-2002 JP 2000280990  
PF 28-Jul-2000 JP 2000280990  
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAKORU  
PI SAITO,  
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,  
PI KEIICHI NAGAI,TETSUJI OTSUKI  
PC  
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC  
10,  
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT CDS Location/Qualifiers  
1..1772  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 473 a 396 c 419 g 484 t  
ORIGIN

Query Match 18.4%; Score 499.8; DB 6; Length 1772;  
Best Local Similarity 68.5%; Pred. No. 1.7e-133;  
Matches 842; Conservative 1; Mismatches 333; Indels 53; Gaps 9;

QY 257 GAATGTGACATATCAGAAATCCAGAGATACGAGTTAAAGGGGATCTCAACAAAGAG 316  
DB 261 GGGATGTGAATATTGGACATC---CAGATCTGGCTTAATGGGATTTCCGCAAGAG 317  
QY 317 ATGTGTCTGTGATCAATGAAGATGTGACTGTGATGACCATGGAGCTTCTGTGACAGA 376  
DB 318 ATGTGTCTGTGATCAATGAAGATGTGACTGTGATGACCATGGAGCTTCTGTGACAGA 377  
QY 377 TACAGTTCCCTGGTCTTATGATGATTAATAATTAAGACTGAATTTAGACTCAAGAG 436  
DB 378 TCCAAATCCAGGATATGATGATGAATAATTAACCTGAGTTGTCTATCAACACAG 437  
QY 437 CCAAGTCACTCCAGCTCAGCTCAGCTCCAGGGGAGCTGACTACAGCTTCCCAAGAACCC 496  
DB 438 CCAAGTCACTCCAGCTCAGCTCAGCTCCAGGGGAGCTGACTACAGCTTCCCAAGAACCC 497  
QY 497 TACCAAGAGAGAAATGG---TTCAGAGACACAGACACTGGTGAACCTCCATATATACA 553  
DB 498 TTACCAAGAGAGACATGGCCACAGAGACAGACACTGGGAGACCTCCCTGATATATA 557  
QY 554 ATGGAACAAATTTCCAC-----ATGGCTGATGAAA 586  
DB 558 ATCTAACACAAATATCCACATTTGCCAATGAGTTACGGGAGCTTGAATTTGGCCAAATGACT 617  
QY 587 TTAGGACTCTGGAGAAACGATCGAAGCTGCTATCCATTTGAGTGGAGTCTGTG 646  
DB 618 TACGGAGCTCTGGAGCAACCATCGAATGAGCATCTACATCGAGAGGAGATCTGTG 677  
QY 647 GGTGACCTCTGGACCTTATGATGTGTCTTAATCTTAATGATTTCCGTGAAGAAA 706  
DB 678 GGTGACCTCTGGACCTTATGATGTGTCTTAATCTTAATGATTTCCGTGAAGAAA 737  
QY 707 AGAAGTATGAGATTGAGCTTATTAACACTGAGCCAACTGCTCCAGAGAGTTGGCAA 766  
DB 738 AGAAGTATGAGATTGAGCTTATTAACACTGAGCCAACTGCTCCAGAGAGTTGGCAA 797  
QY 767 ATGAGAGAGAGTGAAGATTCGCTCTGAGAGAAATATCTACATGAGAGAGAACGAT 826  
DB 798 ATGAGAGAGAGTGAAGATTCGCTCTGAGAGAAATATCTACATGAGAGAGAACGAT 857  
QY 827 ATGAGAGAGAGTGAAGATTCGCTCTGAGAGAAATATCTACATGAGAGAGAACGAT 886  
DB 858 ATGAGAGAGAGTGAAGATTCGCTCTGAGAGAAATATCTACATGAGAGAGAACGAT 917  
QY 887 CGCTCTGAGACTGCCACTTTAAAGC-----TGCTTCAATTCTGACTTTGG 935  
DB 918 AACCTTTGGGTTGTGCTTTGCAATGATGATCAACACCTTAATTTTGAAGCTTGT 977  
QY 936 TATTTCCCTTTKGGAGAAATCATGTGATGATGATGATGATGATGATGATGATGATGAT 995  
DB 978 GTTTGTCTTTTAAAGAACTATGAGCTGTGTCACTGAGCTGTT--TTGAGGTTTGT 1035  
QY 996 CCAAGCAGCTGAGAGAAAGATTCAGTTTCTGGGAGATTAATTAATCAACAGGGGATTC 1055  
DB 1036 CCAAGCAGCTGAGAGAAAGATTCAGTTTCTGGGAGATTAATTAATCAACAGGGGATTC 1094  
QY 1056 GACTGTAACTCATGCTCAATGAATGAAATGCTCCATTTTAATCCCTGAGTTTCAAGG---ATCG 1112  
DB 1095 AACTGGAGACCTGCACTGAATTAACAGAGCATGCTGATGCTCTGATTTAAAGCAACAG 1154  
QY 1113 GATCTCCACCTCCAGAGACT--TCAATGATGCTGTTGAAGTCACTGCTGTTT--CATTA 1169  
DB 1155 AGTTACCAACCCAGAGACTGTTAATGATGATGATGATGATGATGATGATGATGATGAT 1214  
QY 1170 CATTAGGAATGTTAGTGTGATGTTTG 1198  
DB 1215 CACTAGGAATTTGAGTGGGCTCTG 1243

RESULT 14  
AK027334 1772 bp mRNA linear PRI 01-AUG-2002  
LOCUS

DEFINITION Homo sapiens cDNA FLJ14428 fis, clone HEMBA1006293.  
ACCESSION AK027334  
VERSION AK027334.1 GI:14041941  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Iisogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Makamatsu,A., Nakamura,Y., Nagahara,K., Masuno,Y. and Oshima,A.  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1772)  
AUTHORS Iisogai,T. and Otsuki,T.  
TITLE Direct Submission  
JOURNAL Submitted (10-MAY-2001) Takao Iisogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yama, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert construction; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.  
FEATURES  
source  
Location/Qualifiers  
1..1772  
/organism="Homo sapiens"  
/mol\_type="mrna"  
/db\_xref="taxon:9606"  
/clone="HEMBA1006293"  
/tissue\_type="whole embryo, mainly head"  
/clone\_lib="HEMBA1"  
/dev\_stage="embryo, 10 weeks"  
/note="Cloning vector: pME185FL3"  
44..949  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB55044.1"  
/db\_xref="GI:14041942"  
/translation="MESHLPFDVLLLLLLLRSESEVRAEYGNANLPCTTPAA  
PENLVPVCKGKQAFVEGCVNVLRTDERDVNTSRVWLNDFRGGVSLTENVTL  
ADSGLYCQRIQIPGINDEKFMKLVIKPAKVPAPTLQRDFPAAPRMLTTRGHPLA  
ETQTLGSLPDINLTQISTLANELRDLRLANDLSGATIRIGIYIAGICAGIALALI  
FGALLPKWYSHSKERIKONLSLISLANLPSGLANAVAEGRSEBNITYIEENVYEE  
PNEYCYVSRQPSQPLGCRFAMP"  
BASE COUNT 473 a 396 c 419 g 484 t  
ORIGIN  
Query Match 18.4%; Score 499.8; DB 9; Length 1772;  
Best Local Similarity 68.5%; Pred. No. 1.7e-133;  
Matches 842; Conservative 1; Mismatches 333; Indels 53; Gaps 9;  
QY 17 CGGATTTCCCTCCCAAGTACTGATGTTTTCAGGCTTACCTCAACTGTGCTCTGCTG 76  
DB 21 CTGACTTTTCTTGTGCAAGCTCCATGTTTTCACATCTTCCCTTGACTGTCTCTGCTG 80  
QY 77 TGCTGCACTACTACTTGAAGTCACTGGAAGTGAATGATTAAGTGAAGTTGATTA 136  
DB 81 TGTCTGTCTACTACTTCAAGGTCCTCAAGAGTGAATTAACAGAGCGAGGTGGTCA 140  
QY 137 ATGCTATCTGCTCCCTGCACTTCACTTCACTTCACTTGGGACACTTGTGCTATGCT 196  
DB 141 ATGCTATCTGCTCCCTGCTTCACTCACTCACTGCGCCGCGGAGGACCTGCTGCT 200  
QY 197 GGGGCAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 256  
DB 201 GGGGCAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 260

QY 257 GAATGTGACATATCAGAAATCCAGACATACCAGCTAAAGGCGATCTCAACAAAGG 316  
DB 261 GGAATGTGATTAATGACATC---CAGATACGTGCTAAATGGGATTTCCCAAGGAG 317  
QY 317 ATGTGCTCTGATCAATAAGATGTGACCTCTGGATGACCAATGGACCTACTGTCGAGA 376  
DB 318 ATGTGCTCTGATCAATAAGATGTGACCTCTGGATGACCAATGGACCTACTGTCGAGA 377  
QY 377 TACAGTCCCTGCTGCTTATGAATGATTAATAAATTAAGACTGAATTAAGCAATCAAGCAG 436  
DB 378 TCCAAATCCAGGCAATTAAGATGATGAATAAATTAAGCTGAAGTTGGTCATCAACCG 437  
QY 437 CCAAGGTCACTCCAGTCAAGTATGCTCCATGGGACCTCTACTACAGCTTCCAAAGACC 496  
DB 438 CCAAGGTCACTCCAGTCAAGTATGCTCCATGGGACCTCTACTACAGCTTCCAAAGATTC 497  
QY 497 TAACACGAGAGAAATGG---TTGAGAGACACAGACACTGTGTGACCTCCATTAATTA 553  
DB 498 TTACACGAGAGAGAAATGG---TTGAGAGACACAGACACTGTGTGACCTCCATTAATTA 557  
QY 554 ATGAAACAAAATTTCCAC-----ATGGGCTGATGAA 586  
DB 558 ATCTAACACAAATATCCACATTTGGCCAAATGATTAAGGACTCTAGATTTGGCCAATGACT 61  
QY 587 TTAAGGACTCTGAGAAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 646  
DB 618 TACGGGACTCTGAGAAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 677  
QY 647 GGTGACCTGACCTATATCATTTGCTTATCTTAAATGATTAATCTTCTTAAGAA 706  
DB 678 GGTGACCTGACCTATATCATTTGCTTATCTTAAATGATTAATCTTCTTAAAGAA 737  
QY 707 AAGATTTCAAGTTTGAAGCTTATTAATCACTGACCACTTGTCTCAGAGGCTTGCA 766  
DB 738 AAGATTTCAAGTTTGAAGCTTATTAATCACTGACCACTTGTCTCAGAGGCTTGCA 797  
QY 767 ATGAGGAGCACTGAGATGCTGCTGAGAAATTAATCAACCACTGAGAGAAAGCTAT 826  
DB 798 ATGAGGAGCACTGAGATGCTGCTGAGAAATTAATCAACCACTGAGAGAAAGCTAT 857  
QY 827 ATGAGGAGCACTGAGATGCTGCTGAGAAATTAATCAACCACTGAGAGAAAGCTAT 886  
DB 858 ATGAGGAGCACTGAGATGCTGCTGAGAAATTAATCAACCACTGAGAGAAAGCTAT 917  
QY 887 CCGCTCTGAGCTGCACTTTAAAGC-----TCGCTTCATTTCTGACTTTG 935  
DB 918 AACCTTTGGGTTGTGCTTGGCAATGCAATGATCAACCACTTATTTTGAAGCTTGT 977  
QY 936 TATTTCCCTTTTGAAGAACTATGATATGCACTGAGCAACCTCATTTGAGGTTCTGA 995  
DB 978 GTTTGTCTTTTCAAGAACTATGATATGCACTGAGCAACCTCATTTTGAAGCTTGT 1035  
QY 996 CCACAGCACTGAGAAAGAGTTTCCAGTTTCTGAGGATTAATTAATCAAGAGGATTC 1055  
DB 1036 CCACAGCACTGAGAAAGAGTTTCCAGTTTCTGAGGATTAATTAATCAAGAGGATTC 1094  
QY 1056 GACTGATCACTGATCACTTGAATGCTTCACTTTTATCTGAGTTTCAGG---ATCG 1112  
DB 1095 AACTGAGCACTGATCACTTGAATGCTTCACTTTTATCTGAGTTTCAGG---ATCG 1154  
QY 1113 GATTCCTCCACTCAGAGACT--TCATTCATGAGGCTTGAAGCTCAGTCGCTT--CAT 1169  
DB 1155 AATTACCACTCAGAGACTTGAATGCTTGAATGCTTGAAGCTCAGTCGCTT--CAT 1214  
QY 1170 CATTAAGATGCTGATGATGCTTTG 1198  
DB 1215 CACTAGGAATCTTGAACGAGGCTCTCTG 1243  
RESULT 15  
ARI43568 2236 bp DNA linear PAT 08-AUG-2001  
LOCUS ARI43568  
DEFINITION Sequence 23 from patent US 6204371.

ACCESSION ARI43568  
VERSION ARI43568.1 GI:15104854  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 2236)  
TITLE Levinson, D. Adam.  
COMPOSITIONS AND METHODS FOR THE TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
JOURNAL Patent: US 6204371-A 23 20-MAR-2001;  
FEATURES Location/Qualifiers  
source 1..2236  
BASE COUNT 586 a 518 c 525 g 606 t 1 others  
ORIGIN  
Query Match 18.4% Score 499.8; DB 6; Length 2236;  
Best Local Similarity 68.5%; Pred. No. 1.7e-133;  
Matches 842; Conservative 1; Mismatches 333; Indels 53; Gaps 9;  
QY 17 CGGATTTCCCTCCCAAGTACTCATGTTTTCAGGCTTACCTCACTGTGTCTGCTG 76  
DB 19 CTGACTTTCTTCTGCAAGCTCCATGTTTTCATCTTCCCTTGAATGTGTCTGCTG 78  
QY 77 TGCGCACTACTACTGCAAGGTCATTTGAAGATGTTATAGGTTGAGGTTGATAAA 136  
DB 79 TGCTGCTCTACTACTTTCAGAGGCTCTCAAGATGTAATACAGACCGAGGTGCTGCA 138  
QY 137 ATGCTATCTGCTCCGTCAGATTACCTTACCTTACCTTGGGACACTTGTGCTATGCT 196  
DB 139 ATGCTATCTGCTCCGCTTCTTACACCCAGCCGCCCGGAGAACCTGTCGCTGCT 198  
QY 197 GGGGCAAGGATTTCTGCTCTTGTGTCACAGTGTACCAATGATGTTGCTCAAGTGA 256  
DB 199 GGGGCAAGGAGCTGCTGCTGTTTGAATGTGCAAGTGTGCTCAAGTGA 258  
QY 257 GAAATGTACATATCAGAAATCCAGCAGTACCACTAAAGGGGAGTCTCAACAAAGAG 316  
DB 259 GGGATGTAAATTTATGACATC--CAATATGCTAAATGAGGATTTCCGCAAGAG 315  
QY 317 ATGTGCTCTGATCATTAAGATGTGACTGTGATGACCATGGACCTTACTGTGACAGA 376  
DB 316 ATGTGCTCTGATCATTAAGATGTGACTGTGATGACCATGGACCTTACTGTGACAGA 375  
QY 377 TACAGTTCCCTGCTTATGATGTAATAAATTGAAGTGAATTAACATCAAGCAG 436  
DB 376 TCCAAATCCCAAGGCAATGATGTAATAAATTGAAGTGAATTAACATCAAGCAG 435  
QY 437 CCAAGGTACTCCAGCTCAGACTGCGCATGGGAGCTTACTACAGCTTCTCCAAAGACC 496  
DB 436 CCAAGGTACTCCAGCTCAGACTGCGCATGGGAGCTTACTACAGCTTCTCCAAAGATGC 495  
QY 497 TAACACGAGAGAAATG---TTGAGACACAGACACTGCTGACCTCCATTAATAACA 553  
DB 496 TTACACGAGAGGAGATGCGCCAGACAGACACACTGCGGAGCTCCCTGATATA 555  
QY 554 ATGAAACAAAATTTCCAC-----ATGGGCTGATGAA 586  
DB 556 ATCTAACACAAATATCCACATTTGGCCAAATGAGTTACGGGAGCTTGAATTTGGCAATGACT 615  
QY 587 TTAAGGACTCTGAGAAAGATCAGACATGCTATCCACATTTGAGAGTCTGCTG 646  
DB 616 TAGGGACTCTGAGCAACATCAGAAATGAGCATCTACATGAGACAGAGATCTGTGCTG 675  
QY 647 GGTGACCTGCGCATTAATGATGTTGCTTAACTTAATGATGATTTCTGTAGAAAA 706  
DB 676 GGTGCTGCTGCTTATCTTGTGCTTAAATTTTCAATGATTTCTCATAGCAAG 735  
QY 707 AGAAGTTATGAGTTTGAAGCTTATTAACATGCGCAACTTGTGCTCAGAGGCTTGCA 766  
DB 736 AGAAGTTATGAGTTTGAAGCTTATTAACATGCGCAACTTGTGCTCAGAGATTTGCA 795

QY 767 ATGAGAGGAGCTGAGGATTCGCTCTGAGAAAAATATCTACACCATCAGAGAGACGTAT 826  
DB 796 ATGAGATGAGAGAGGAAATTCGCTCAGAAAAACATCTATACATTTGAGAGAGACGTAT 855  
QY 827 ATGAGTGAAGATTTCAATGATGATCTACTGCTACGTCACAGCCAGACCATCTGAC 886  
DB 856 ATGAGTGAAGAGACCCCAATGATGATTTGCTATGTCAGACAGCAGCAACCTTCAC 915  
QY 887 GCGCTGAGACTGCGACTTTTAAAGG-----TGCCCTTAATTTCTGACTTTG 935  
DB 916 AACCTTGGGTTGTGCTTGGCAATGCAATGATCAACCACTTATTTTGAAGCTTGT 975  
QY 936 TATTTCCCTTTTGTGAAAAACTATGATGATGATGATGATGATGATGATGATGATGAT 995  
DB 976 GTTTTGTCTTTTTCAGAAACTATGATGATGATGATGATGATGATGATGATGATGAT 1033  
QY 996 CCACAGCCACTGAGAAAAAGATTCAGTTTCTGAGGATTAATTAATCTCAAGGAGATTC 1055  
DB 1034 CCACCTGCTATGAGACAGAGTTTTCCTTTCAGAGATTA--TGACTCACATGAGGAAATTTG 1092  
QY 1056 GACTGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1112  
DB 1093 AACTGGAGCTGCACTGAATTAACAGGCAATGATGATGATGATGATGATGATGATGAT 1152  
QY 1113 GATCTCCCACTCAGAGACT--TCAATCATGCGTGTGAGCTCACTGCTTT--CAT 1169  
DB 1153 AGTTACCAACCCAGAGACTCTTATCATGATGATGATGATGATGATGATGATGATGAT 1212  
QY 1170 CATTAGAAATGTTAGTGTGATGCTTTG 1198  
DB 1213 CACTAGAAATCTTGAAGTGTGCTGCTG 1241

Search completed: November 21, 2003, 23:34:26  
Job time : 10139.7 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 11:32:52 ; Search time 747.069 Seconds  
(without alignments)  
9792.239 Million cell updates/sec

Title: US-10-004-633-8  
Perfect score: 2710  
Sequence: 1 ntgcgaccacgcgcgcgcga.....aaaaaaagcgccgcgc 2710

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N Geneseq 19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneeqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2706.6	99.9	2710	21	AAAS1898
2	2706.6	99.9	2710	22	AAI70263
3	2706.6	99.9	2710	22	AAF82609
4	2706.6	99.9	2710	22	AAC90986
5	2706.6	99.9	2710	22	AAF23458
6	2706.6	99.9	2710	24	AB53316
7	2705	99.8	2710	25	ABQ77037
8	2684.6	99.1	2712	17	AAT38265

9	2624.6	96.8	2725	25	AB268328
10	844.4	31.2	862	25	AB268329
11	843	31.1	843	21	AAAS1918
12	843	31.1	843	22	AAI70255
13	843	31.1	843	22	AAF82615
14	843	31.1	843	22	AAC90992
15	843	31.1	843	22	AAF23464
16	843	31.1	843	25	ABQ77038
17	499.8	18.4	1772	22	AAH15441
18	499.8	18.4	2236	21	AAAS1904
19	499.8	18.4	2236	22	AAI70254
20	499.8	18.4	2236	22	AAF82616
21	499.8	18.4	2236	22	AAC90993
22	499.8	18.4	2236	22	AAF23465
23	499.8	18.4	2236	24	AB533132
24	499.8	18.4	2236	25	ABQ77040
25	499.8	18.4	1756	22	AAAS2714
26	499	18.4	1475	24	ABL90805
27	497.8	18.4	1800	22	AAAS2478
28	497.2	18.3	2237	17	AAI78267
29	485.8	17.9	1116	25	AB268338
30	485.8	17.9	1116	25	AB268339
31	456	16.8	903	21	AAAS1905
32	456	16.8	903	22	AAI70281
33	456	16.8	903	22	AAF82629
34	456	16.8	903	22	AAC92155
35	456	16.8	903	22	AAF23478
36	456	16.8	903	25	ABQ77041
37	384	14.2	1326	25	ABT33381
38	336.4	12.4	769	22	AAH04055
39	328.8	12.1	2091	20	AAAS35717
40	289	10.7	555	20	AAAS35716
41	263.4	9.7	529	20	AAV87937
42	231.6	8.5	1203	21	AAAS8799
43	219.8	8.1	438	20	AAAS1096
44	84.2	3.1	1079	25	AB268335
45	84.2	3.1	1080	25	AB268332

## ALIGNMENTS

RESULT 1	AAAS1898
ID	AAAS1898 standard; DNA; 2710 BP.
AC	AAAS1898;
XX	XX
DT	31-OCT-2000 (first entry)
DE	Murine T helper cell differentially expressed gene 200.
XX	XX
KW	T helper cell; differential expression; 200 gene; immunomodulator; anti-inflammatory; anti-arthritic; antibacterial; immunosuppressive; chryomimetic; anti-chyroid; anti-asthmatic; anti-allergic; antiviral; protozoacide; lymphocyte; modulator; gene therapy; ss.
KW	XX
OS	Mus sp.
XX	XX
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	40..885
FT	/*tag= a
FT	40..99
FT	/*tag= b
FT	100..882
FT	/*tag= c
PD	US6084083-A.
XX	XX
PD	04-JUN-2000.
XX	XX
PF	28-MAR-1997; 97US-0829525.
XX	XX

Nucleotide sequenc  
Nucleotide sequenc  
T helper cell diff  
Mouse 200 gene, pr  
Murine TH1 specifi  
DNA associated wit  
Murine TH1-associ  
Human cDNA sequenc  
Human T helper cel  
Human 200 gene, pr  
Human TH1 specific  
Human 200 gene nuc  
Human 200 gene. H  
DNA encoding human  
Human TH1-associat  
Human cDNA encodin  
Human polynucleoti  
Human cDNA encodin  
200 gene different  
Nucleotide sequenc  
Nucleotide sequenc  
Human T helper cel  
Human 200 gene, pr  
Human TH1 specific  
Human 200 gene nuc  
DNA associated wit  
Human TH1-associat  
NOVX DNA sequenc  
Human cDNA clone (c  
cDNA encoding a pr  
EST clone FG708.  
Human SECK cDNA C1  
Human secreted pro  
Nucleotide sequenc  
Nucleotide sequenc

PR	01-MAR-1996;	96US-0609583.
PR	03-MAR-1995;	95US-0398633.
PR	07-JUN-1995;	95US-0487748.
XX		
PA	(MILL-) MILLENNIUM PHARM INC.	
XX		
PI	Levinson DA;	
DR	WPI; 2000-464385/40.	
DR	P-PSDB; AAV97056.	
XX		
PT	New isolated human 200 gene products or polypeptides, useful for	
PT	treating and diagnosing immune disorders, especially T helper	
XX	lymphocyte-related disorders	
PS		
XX	Example; Fig 17A-D; 107pp; English.	
CC		
CC	Genes which are differentially expressed within and among T helper (TH)	
CC	cells and TH cell subpopulations, e.g. TH0, TH1 and TH2 subpopulations,	
CC	can be used diagnostically or as targets for therapeutic intervention.	
CC	The polypeptides are useful for treating and diagnosing of immune	
CC	disorders, especially T lymphocyte-related disorders. These disorders	
CC	include chronic inflammatory diseases and disorders (e.g. Crohn's	
CC	disease), reactive arthritis, Lyme disease, Hashimoto's thyroiditis or	
CC	Grave's disease), or atopic conditions (e.g. asthma and allergy,	
CC	including allergic rhinitis or food allergies). Also included are	
CC	certain pathogen susceptibilities (e.g. leishmaniasis), and viral (e.g.	
CC	HIV) or bacterial (e.g. tuberculosis or leptomatous leprosy) infections.	
XX		
SQ	Sequence 2710 BP; 789 A; 592 C; 644 G; 678 T; 7 other;	

Query Match	Similarity	99.9%	Score	2706.6	DB	21	Length	2710
Beet	Local	Similarity	100.0%	Pred.	No.	0		
Matches	2709	Conservative	0	Mismatches	0	Indels	0	Gaps
QY	2	GTGACCCACGCGGTCGCGATTTTCCCTCCCAAGTACTATGTTTTCAGGTCTTACCTCA	61					
Db	2	GTGACCCACGCGGTCGCGATTTTCCCTCCCAAGTACTATGTTTTCAGGTCTTACCTCA	61					
QY	62	ACTGTGCTCTGTGCTGCTGCAACTACTTTCGCAAGGCACTTGAAGAAGTGTATTAAG	121					
Db	62	ACTGTGCTCTGTGCTGCTGCAACTACTTTCGCAAGGCACTTGAAGAAGTGTATTAAG	121					
QY	122	TTGAGGTGTGTAATAATGCTTCTTCCCTCGCAGTTACCTCTACTCATCTGGGAGAC	181					
Db	122	TTGAGGTGTGTAATAATGCTTCTTCCCTCGCAGTTACCTCTACTCATCTGGGAGAC	181					
QY	182	TTGTGCTCTATGTGCTGTGGGCAAGGATTTGTTCCTTGTGCACAGTGTACCAATGAGTTC	241					
Db	182	TTGTGCTCTATGTGCTGTGGGCAAGGATTTGTTCCTTGTGCACAGTGTACCAATGAGTTC	241					
QY	242	TCAGAACTGATCAAAAGAAATGTGACATATCAGAAATCCAGCAGATACCGATTAAGGGCG	301					
Db	242	TCAGAACTGATCAAAAGAAATGTGACATATCAGAAATCCAGCAGATATCCAGATTAAGGGCG	301					
QY	302	ATCTCAACAAAGAGATGTGTCTTGATATATAAGATGTGACTCTGGATATACATGGGA	361					
Db	302	ATCTCAACAAAGAGATGTGTCTTGATATATAAGATGTGACTCTGGATATACATGGGA	361					
QY	362	CCTACTGCTGCAGGATATACAGTTCCTCGTCTTATGATGATATAAAAATTAGAACTGAAT	421					
Db	362	CCTACTGCTGCAGGATATACAGTTCCTCGTCTTATGATGATATAAAAATTAGAACTGAAT	421					
QY	422	TGACATCAAAACAGCCAGGCTCATCTCCAGCTCAGATCCCATGGGACTCTTACTACAG	481					
Db	422	TGACATCAAAACAGCCAGGCTCATCTCCAGCTCAGATCCCATGGGACTCTTACTACAG	481					
QY	482	CTTCTCCAAAGACCTTAACCAAGAGAGAAATGTTTACAGACACAGACACTGTGACCC	541					
Db	482	CTTCTCCAAAGACCTTAACCAAGAGAGAAATGTTTACAGACACAGACACTGTGACCC	541					
QY	542	TCGATTAATACATGGACAAAAATTTCCACATGGGCTGATGAATTAAGACTCTGGAG	601					
Db	542	TCGATTAATACATGGACAAAAATTTCCACATGGGCTGATGAATTAAGACTCTGGAG	601					

Db	542	TCCTATTAAACAATGGAACAAAAATTTCCACTGAGGCTGATGAAATTAAGACCTCTGGAG	601
Qy	602	AAACGATCAGAAGCTGTATCCACTTGGAGTGGGAGTCTCTGTGGGTGACCTTGCAAC	661
Db	602	AAACGATCAGAAGCTGTATCCACTTGGAGTGGGAGTCTCTGTGGGTGACCTTGCAAC	661
Qy	662	TTATCATTTGGTGTCTTAATCCTTAAATGATTCCTGTAAAGAAAAGACGTTATCGAGTT	721
Db	662	TTATCATTTGGTGTCTTAATCCTTAAATGATTCCTGTAAAGAAAAGACGTTATCGAGTT	721
Qy	722	TGAGCCTTATTAACACTGAGCAACTTGGCTCAGAGAGGTTGGCAAAATGCAAGAGCACTCA	781
Db	722	TGAGCCTTATTAACACTGAGCAACTTGGCTCAGAGAGGTTGGCAAAATGCAAGAGCACTCA	781
Qy	782	GGATTGCTCTGAGGAAAATATCTACACCATCGAGAGAAAGTATATGAATGAGAAATT	841
Db	782	GGATTGCTCTGAGGAAAATATCTACACCATCGAGAGAAAGTATATGAATGAGAAATT	841
Qy	842	CAAAATGAGTACTACTGCTACGCTCAACAGCCAGAGCCATCTGACCCGCTCTGGACTGCC	901
Db	842	CAAAATGAGTACTACTGCTACGCTCAACAGCCAGAGCCATCTGACCCGCTCTGGACTGCC	901
Qy	902	ACCTTTAAAGGCTGCGCCTTCATTTCTGACCTTGGTATTTCCCTTTKTGAAAAATCTATGTG	961
Db	902	ACCTTTAAAGGCTGCGCCTTCATTTCTGACCTTGGTATTTCCCTTTKTGAAAAATCTATGTG	961
Qy	962	ATATGTCACTTTGGCAACCTCATTTGGAGTTCTGACCAAGCCACTGAGAAAAAGAGTTCCA	1021
Db	962	ATATGTCACTTTGGCAACCTCATTTGGAGTTCTGACCAAGCCACTGAGAAAAAGAGTTCCA	1021
Qy	1022	GTTTTCTGGGGATTAATTAATCAACAAAGGGATTCGACTGTAACTCAATGATTAAT	1081
Db	1022	GTTTTCTGGGGATTAATTAATCAACAAAGGGATTCGACTGTAACTCAATGATTAAT	1081
Qy	1082	GCTCCATTTTATCCCTGAGTTTCAGGGATCGAGTCTCCCACTCCAGAGACTTCAATCATGT	1141
Db	1082	GCTCCATTTTATCCCTGAGTTTCAGGGATCGAGTCTCCCACTCCAGAGACTTCAATCATGT	1141
Qy	1142	CGTGTGAAGCTCACTCGTGTCTTATCATATTTAGAAATGTTAGTGTGATCTTTGAGA	1201
Db	1142	CGTGTGAAGCTCACTCGTGTCTTATCATATTTAGAAATGTTAGTGTGATCTTTGAGA	1201
Qy	1202	CATAGAGTTTGTGTGATATATCCGAAAAGCTCTGAAACAGGTAGGGGGAATTAAGGCTTAA	1261
Db	1202	CATAGAGTTTGTGTGATATATCCGAAAAGCTCTGAAACAGGTAGGGGGAATTAAGGCTTAA	1261
Qy	1262	GATAGGAAGTGGCGGTCTTTGTTGATGTGGAATAATCTTAAAGAAAGTTGGAGCTTTTCT	1321
Db	1262	GATAGGAAGTGGCGGTCTTTGTTGATGTGGAATAATCTTAAAGAAAGTTGGAGCTTTTCT	1321
Qy	1322	AGAGATTTCTGACCTTGAAAGATTTAAGAAAAAGCCAGGTGGCATATGCTTAAACGATAT	1381
Db	1322	AGAGATTTCTGACCTTGAAAGATTTAAGAAAAAGCCAGGTGGCATATGCTTAAACGATAT	1381
Qy	1382	AACTTGGGAACCTTAGAGCAGAGAGGTGATTAAGTTCAAGGTCAAGGTCAAGGACTATCTGTGA	1441
Db	1382	AACTTGGGAACCTTAGAGCAGAGAGGTGATTAAGTTCAAGGTCAAGGTCAAGGACTATCTGTGA	1441
Qy	1442	AGACTGTCTCMACATCCCAAAGACGAAAAATAACTAAGACAGCAGCAGAGGCTGAGAGATGA	1501
Db	1442	AGACTGTCTCMACATCCCAAAGACGAAAAATAACTAAGACAGCAGCAGAGGCTGAGAGATGA	1501
Qy	1502	GGCTTCGACAGTGAAGTGCATTTGTGTACAAAGCAGAGAAATCTATATTGATCTGTAGACC	1561
Db	1502	GGCTTCGACAGTGAAGTGCATTTGTGTACAAAGCAGAGAAATCTATATTGATCTGTAGACC	1561
Qy	1562	CCACATGAAAAAGCTAGGCTCTGTGATGAGCAATGCTTTGATGACTCAAGATGAGTAGAGGTAA	1621
Db	1562	CCACATGAAAAAGCTAGGCTCTGTGATGAGCAATGCTTTGATGACTCAAGATGAGTAGAGGTAA	1621
Qy	1622	AGGCAACAACAATCCCGGGGCTTGGCGGACAGTACAGCTTAAGGTGCTGAGTTCCAA	1681
Db	1622	AGGCAACAACAATCCCGGGGCTTGGCGGACAGTACAGCTTAAGGTGCTGAGTTCCAA	1681

QY 1682 GTCCACAAAGTCCCTGTCTCAGAGTATGAGTCTGAGTATCTGGCGCATGTCCATGGG 1741  
 DB 1682 GTCCACAAAGTCCCTGTCTCAGAGTATGAGTCTGAGTATCTGGCGCATGTCCATGGG 1741  
 QY 1742 GGTGTCTCTCTCTCTCAGAGAGACATGACATGCTCTGACACACACACACACAC 1801  
 DB 1742 GGTGTCTCTCTCTCTCAGAGAGACATGACATGCTCTGACACACACACACACAC 1801  
 QY 1802 AC 1861  
 DB 1802 AC 1861  
 QY 1862 TCTCTATAACATGTATCTCTACAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
 DB 1862 TCTCTATAACATGTATCTCTACAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
 QY 1922 CATGCGACAGCAGTCCAGTAATTTATTCAGACACTCAGAAAGCTGGAGCAAGCGTGA 1981  
 DB 1922 CATGCGACAGCAGTCCAGTAATTTATTCAGACACTCAGAAAGCTGGAGCAAGCGTGA 1981  
 QY 1982 GAGTTCAGGAGCAGTGTGCCCCAACACTGCGACACTCTCTTACAGAAAGGTTACC 2041  
 DB 1982 GAGTTCAGGAGCAGTGTGCCCCAACACTGCGACACTCTCTTACAGAAAGGTTACC 2041  
 QY 2042 CGCAAGCAGCCTGTCTGTCTGTAAAGAAACCTGCGAAAGGCAAACTTTGACTGTGTG 2101  
 DB 2042 CGCAAGCAGCCTGTCTGTCTGTAAAGAAACCTGCGAAAGGCAAACTTTGACTGTGTG 2101  
 QY 2102 TGTCTAAGGGGAACTGACTCAGACAACTTCTCATCTCTGAGAAAGCTGAGCTGTTC 2161  
 DB 2102 TGTCTAAGGGGAACTGACTCAGACAACTTCTCATCTCTGAGAAAGCTGAGCTGTTC 2161  
 QY 2162 TGACAGAAAGAACACCGGTGACTGGGACATACGAAAGGAGACCTCTGAGCAATCTTA 2221  
 DB 2162 TGACAGAAAGAACACCGGTGACTGGGACATACGAAAGGAGACCTCTGAGCAATCTTA 2221  
 QY 2222 TAGTCAGCAAAATATTTCTTTGGGAGACAGTGTGACCAAAATTTGATTTCCAAAGCGGTG 2281  
 DB 2222 TAGTCAGCAAAATATTTCTTTGGGAGACAGTGTGACCAAAATTTGATTTCCAAAGCGGTG 2281  
 QY 2282 ACCTTAAGTTTCACTGTGGCTTAAGCTGCTGCCAGTGTGATCTGTGTGCTGCC 2341  
 DB 2282 ACCTTAAGTTTCACTGTGGCTTAAGCTGCTGCCAGTGTGATCTGTGTGCTGCC 2341  
 QY 2342 ATCTATAACAGATCAAAATTAATAGACCCGAGGAAATTTAAGTAGACAGAAAGT 2401  
 DB 2342 ATCTATAACAGATCAAAATTAATAGACCCGAGGAAATTTAAGTAGACAGAAAGT 2401  
 QY 2402 AGCTTTGTTCAAAAGATTTTTCATTTGGGAGCACTGTGTACATCAGAGACATCTGT 2461  
 DB 2402 AGCTTTGTTCAAAAGATTTTTCATTTGGGAGCACTGTGTACATCAGAGACATCTGT 2461  
 QY 2462 TAGTAGAGACACCAAAACCTGTGTGACCGTTTTCATGTATGAATTTGTTTAAAGT 2521  
 DB 2462 TAGTAGAGACACCAAAACCTGTGTGACCGTTTTCATGTATGAATTTGTTTAAAGT 2521  
 QY 2522 TGCTCTAGCTAGCTGTGAGAGTCTGTGCTTTCTAGTGGGTATGAAAGGAGCAATC 2581  
 DB 2522 TGCTCTAGCTAGCTGTGAGAGTCTGTGCTTTCTAGTGGGTATGAAAGGAGCAATC 2581  
 QY 2582 TAAACAAATTCATTAAGATTAACAGCTCTCATGCAAGAGGAAACATATCTCAATGTT 2641  
 DB 2582 TAAACAAATTCATTAAGATTAACAGCTCTCATGCAAGAGGAAACATATCTCAATGTT 2641  
 QY 2642 TTTAAAGTATAAACTGTACTGGCAAGTACTTTAGCATTAATAAAAAAAAAAAAAAAG 2701  
 DB 2642 TTTAAAGTATAAACTGTACTGGCAAGTACTTTAGCATTAATAAAAAAAAAAAAAAAG 2701  
 QY 2702 GGGCGCGC 2710  
 DB 2702 GGGCGCGC 2710

RESULT 2  
 AAI70263  
 ID AAI70263 standard; cDNA, 2710 BP.  
 XX  
 AC AAI70263;  
 XX  
 DT 07-JAN-2002 (first entry)  
 XX  
 DE Mouse 200 gene, preferentially expressed in TH1 cells.  
 XX  
 KW Mouse; 200 gene; T helper; T lymphocyte; T cell; TH1;  
 KW receptor; differential expression; immune disorder; psoriasis;  
 KW multiple sclerosis; insulin-dependent diabetes; antidiabetic;  
 KW antipsoriatic; diagnosis; therapy; se.  
 XX  
 OS Mus musculus.  
 XX  
 FH  
 FT  
 FT CDS Location/Qualifiers  
 FT 40..885  
 FT /\*tag= a  
 XX  
 XX US6288218-B1.  
 XX  
 PD 11-SEP-2001.  
 XX  
 PF 25-SEP-1997; 97US-0937399.  
 XX  
 XX 01-MAR-1996; 96US-0609583.  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 XX  
 PA (LEVI/) LEVINSON D A.  
 XX  
 PI Levinson DA;  
 XX  
 DR WPI, 2001-647189/74.  
 DR P-PSDB; AAM50225.  
 XX  
 PT Detecting 200 gene expression in a sample, useful for treating and  
 PT diagnosing immune disorders, especially T lymphocyte-related disorders,  
 PT comprises detecting the presence of a 200 gene product or an RNA  
 PT encoding the 200 gene product  
 XX  
 PS Claim 2 (n); Fig 17A-D; 108pp; English.  
 CC  
 CC The present sequence is that of cDNA including the complete open  
 CC reading frame of the mouse 200 gene. It encodes a 281-amino acid  
 CC protein (see AAM50225). The cDNA was isolated using the 'T cell  
 CC clone paradigm' and the 'transgenic T cell paradigm' approaches  
 CC to the identification of genes that are differentially expressed  
 CC in helper T cell (TH) subpopulations. The gene is expressed at  
 CC a higher level in fully differentiated TH1 than in TH2 cells. It  
 CC can be used diagnostically or as a target for therapeutic  
 CC intervention for the treatment of immune disorders. A claimed  
 CC method for diagnosing a TH cell subpopulation-related immune  
 CC disorder involves detecting the level of a human 200 gene (see  
 CC AAI70263) product, or an RNA encoding it, so that if the level  
 CC differs from that in a control sample, the disorder is diagnosed.  
 CC The disorder is especially a TH1 cell subpopulation-related immune  
 CC disorder, such as multiple sclerosis, psoriasis or insulin-dependent  
 CC diabetes (claimed). A method for detecting murine 200 gene  
 CC expression is also claimed. In addition to the 200 gene, the  
 CC invention provides other genes that are differentially expressed  
 CC within and among TH cells and TH cell subpopulations and which can be  
 CC used in methods for the diagnosis, prognosis, evaluation and treatment  
 CC of TH cell subpopulation-related disorders, for the identification of  
 CC subjects exhibiting a predisposition to such conditions, for  
 CC monitoring patients undergoing clinical evaluation for the  
 CC treatment of such disorders, and for monitoring the efficacy of  
 CC compounds used in clinical trials. Other immune disorders that can  
 CC be treated/diagnosed include Crohn's disease, reactive arthritis,  
 CC Lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis,





```
Db 1982 GAGTTCAGAGAGACTGTGCCAACCTGCGAGACTCTTTCACACAAGAAAAAGTTACC 2041
Qy 2042 CGCAAGCAGCCTGCTGTCTGTAAAGGAAACCTGCGAAAGCAAACTTGACTGTG 2101
Db 2042 CGCAAGCAGCCTGCTGTCTGTAAAGGAAACCTGCGAAAGCAAACTTGACTGTG 2101
Qy 2102 TGCCTAAGGGAAGTGAAGTCAAGCACTTCTCCATCTCTGAGAAAGTGAAGTCTTTC 2161
Db 2102 TGCCTAAGGGAAGTGAAGTCAAGCACTTCTCCATCTCTGAGAAAGTGAAGTCTTTC 2161
Qy 2162 TCACAGAGAGAAACACCGGTGACTGGGACATACGAAAGGACAGAGCTCTTGCAGCAATCTATA 2221
Db 2162 TCACAGAGAGAAACACCGGTGACTGGGACATACGAAAGGACAGAGCTCTTGCAGCAATCTATA 2221
Qy 2222 TAGTCAGCAAAATATTTCTTTGGGAGAGACGTGTCACAAATTTGATTTCCAGCCGGTGG 2281
Db 2222 TAGTCAGCAAAATATTTCTTTGGGAGAGACGTGTCACAAATTTGATTTCCAGCCGGTGG 2281
Qy 2282 ACCCTAGTTTCATCTGGCTTACAGCTGCTGCCAGTGGCCCTTGAATCTGTGCTGCC 2341
Db 2282 ACCCTAGTTTCATCTGGCTTACAGCTGCTGCCAGTGGCCCTTGAATCTGTGCTGCC 2341
Qy 2342 ATCTATTAACAGATCAATTAATAATAGACCCCGAGTGAATAATTAAGTACAGAAAGT 2401
Db 2342 ATCTATTAACAGATCAATTAATAATAGACCCCGAGTGAATAATTAAGTACAGAAAGT 2401
Qy 2402 AGCTTTGTTCAAGATTTTTTTCATTTGGGAGCAACCTGTGACATCAGAGACATCTGT 2461
Db 2402 AGCTTTGTTCAAGATTTTTTTCATTTGGGAGCAACCTGTGACATCAGAGACATCTGT 2461
Qy 2462 TAGTGAAGACACCAAAACCTGTGTACCTTTTTCATGTATGAATTTTGTGTTAGT 2521
Db 2462 TAGTGAAGACACCAAAACCTGTGTACCTTTTTCATGTATGAATTTTGTGTTAGT 2521
Qy 2522 TGCCTTACGTCAGCTGTGAGAGTCTGCTTTCTTTAGTGGGTATGGAAGGAGACCATC 2581
Db 2522 TGCCTTACGTCAGCTGTGAGAGTCTGCTTTCTTTAGTGGGTATGGAAGGAGACCATC 2581
Qy 2582 TAAACAAATCCATTTAGAGATTAACAGTCTCAGCAGAGGAAAGAACTATCTCAATGTT 2641
Db 2582 TAAACAAATCCATTTAGAGATTAACAGTCTCAGCAGAGGAAAGAACTATCTCAATGTT 2641
Qy 2642 TTAAGTATATAAACTGTACTGTGCAAAAGTACTTTGAGCATATAAAAAAAAAAAG 2701
Db 2642 TTAAGTATATAAACTGTACTGTGCAAAAGTACTTTGAGCATATAAAAAAAAAAAG 2701
Qy 2702 GCGGCGCGC 2710
Db 2702 GCGGCGCGC 2710

RESULT 3
AAF82609
ID AAF82609 standard; cDNA; 2710 BP.
XX
AC AAF82609;
XX
DT 18-JUN-2001 (first entry)
XX
DE Murine TH1 specific 200 gene cDNA.
XX
KW Murine; T helper cell; TH cell; TH1; TH2; immunomodulator;
XX anti-inflammatory; anti-allergic; dermatological; antiviral;
XX antibacterial; T helper lymphocyte modulator; gene therapy;
XX TH specific gene; 200 gene; immune disorder; inflammation;
XX infection; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 40..884
FT /*tag= a
```

```
FT /product= "200 gene product"
XX
PN US6204371-B1.
XX
PD 20-MAR-2001.
XX
PF 01-MAR-1996; 96US-0609583.
XX
PR 03-MAR-1995; 95US-0398633.
XX 07-JUN-1995; 95US-0487748.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Levinson DA;
XX
DR WPI; 2001-272703/28.
XX
PT New murine or human 200 genes and their corresponding polypeptides,
PT useful for treating or diagnosing immune disorders, especially T helper
PT lymphocyte-related disorders, e.g. inflammatory diseases or allergies
PT
PS Claim 1; Fig 17; 109pp; English.
XX
CC The present sequence is expressed at higher levels in T helper 1 (TH1)
CC cells than in T helper 2 (TH2) cells. The invention relates to an
CC isolated nucleic acid molecule, which comprises the full length murine
CC 200 gene or full length human 200 gene nucleotide sequence. The nucleic
CC acids are useful for treating or diagnosing immune disorders. The nucleic
CC acids are useful for treating or diagnosing immune disorders, e.g. inflammatory
CC diseases (e.g. Crohn's disease), multiple sclerosis, Grave's disease,
CC contact dermatitis, psoriasis, asthma and allergies, or certain viral
CC (e.g. HIV) or bacterial (e.g. tuberculosis) infections.
XX
SQ Sequence 2710 BP; 789 A; 592 C; 644 G; 678 T; 7 other;
XX
Query Match 99.9%; Score 2706.6; DB 22; Length 2710;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 GTGACCCACGCGTCGGATTTCCCTCCCAAGTACATGTTTCAAGTCTTACCTTCA 61
Db 2 GTGACCCACGCGTCGGATTTCCCTCCCAAGTACATGTTTCAAGTCTTACCTTCA 61
Qy 62 ACTGTGCTCTGCTGTGTCGCAACTACTACTTTCAGAGTCAATGGAATGTTTAAAG 121
Db 62 ACTGTGCTCTGCTGTGTCGCAACTACTACTTTCAGAGTCAATGGAATGTTTAAAG 121
Qy 122 TTGAGTGTGTAATAATGCTATCTGCTGAGTTTACCTTACCTACATCTGGAGAC 181
Db 122 TTGAGTGTGTAATAATGCTATCTGCTGAGTTTACCTTACCTACATCTGGAGAC 181
Qy 182 TTGTCCTATGTGCTGGGCAAGGATTTCTGTCTTGTGTCAGTGTACAGTGAAGTTC 241
Db 182 TTGTCCTATGTGCTGGGCAAGGATTTCTGTCTTGTGTCAGTGTACAGTGAAGTTC 241
Qy 242 TCAGAACTGATGAAGAATGTGACATATCCAGAAATCCAGAGATTAACAGCTAAAGGGCG 301
Db 242 TCAGAACTGATGAAGAATGTGACATATCCAGAAATCCAGAGATTAACAGCTAAAGGGCG 301
Qy 302 ATCTCAAGAGAGATGTGCTGTGATCATTAAGATGTGACTGTGATGCAATGGGA 361
Db 302 ATCTCAAGAGAGATGTGCTGTGATCATTAAGATGTGACTGTGATGCAATGGGA 361
Qy 362 CCTACTGTGAGATATACAGTTCCCTGTGCTTATGATTAATAAATTAAGAACTGAAT 421
Db 362 CCTACTGTGAGATATACAGTTCCCTGTGCTTATGATTAATAAATTAAGAACTGAAT 421
Qy 422 TAGACATTAAGACAGCCAGGTCACCTCAGCTCAGACTGCCCATGAGGAGCTTACTACAG 481
Db 422 TAGACATTAAGACAGCCAGGTCACCTCAGCTCAGACTGCCCATGAGGAGCTTACTACAG 481
Qy 482 CTCTCCAAAGAACCTTAACACGAGAGAAATGTTCAGAGACACAGACTGTGAGACC 541
```

Db 482 CTTCTCCAGAACCTTAAACCGAGAGAAATGTTTCAGACACAGACACTGAGTACC 541  
Qy 542 TCCATATTAACATGAGAACAAAAATTTCCACATGGGCTGATGAAATTAAGACTCTGAG 601  
Db 542 TCCATATTAACATGAGAACAAAAATTTCCACATGGGCTGATGAAATTAAGACTCTGAG 601  
Qy 602 AAACGATCAGAACTGCTATTCACATTTGAGAGTGGAGTCTCTGCTGGTTGACCTTGGAC 661  
Db 602 AAACGATCAGAACTGCTATTCACATTTGAGAGTGGAGTCTCTGCTGGTTGACCTTGGAC 661  
Qy 662 TTTATCATGCTGCTTATCTTAAATGGAATCTGTAAGAAAAAGATTATGAGTT 721  
Db 662 TTTATCATGCTGCTTATCTTAAATGGAATCTGTAAGAAAAAGATTATGAGTT 721  
Qy 722 TGAACCTTATTAACACTGGCCAACTTGCCTCAGAGAGGTTGGCAAAATGACAGACAGTCA 781  
Db 722 TGAACCTTATTAACACTGGCCAACTTGCCTCAGAGAGGTTGGCAAAATGACAGACAGTCA 781  
Qy 782 GGAATTCCTCTGAGAAAAATATCTACACCATGAGAGAAAGTATATGAAGTGAAGATT 841  
Db 782 GGAATTCCTCTGAGAAAAATATCTACACCATGAGAGAAAGTATATGAAGTGAAGATT 841  
Qy 842 CAAATGAGTACTGCTAGTACGTAAGCAAGAGGAGGATCTGACCGGCTCTGAGACTGCC 901  
Db 842 CAAATGAGTACTGCTAGTACGTAAGCAAGAGGAGGATCTGACCGGCTCTGAGACTGCC 901  
Qy 902 ACTTTTAAAGGCTGGCTTCAATTTCTGACTTTGTAATTTCTCTTTTGTGAAAACTATGTT 961  
Db 902 ACTTTTAAAGGCTGGCTTCAATTTCTGACTTTGTAATTTCTCTTTTGTGAAAACTATGTT 961  
Qy 962 ATATGTCATCTGGCAACCTTCATTTGAGGTTCTGACACAGCCACTGAGAAAAAGTTCCA 1021  
Db 962 ATATGTCATCTGGCAACCTTCATTTGAGGTTCTGACACAGCCACTGAGAAAAAGTTCCA 1021  
Qy 1022 GTTTTCTGGGGATTAATTAATCTACAGAGGGATTCGACTGTAATCTGACTGTAATTAAT 1081  
Db 1022 GTTTTCTGGGGATTAATTAATCTACAGAGGGATTCGACTGTAATCTGACTGTAATTAAT 1081  
Qy 1082 GCTCAATTTTATCCCTGAGTTTCAGAGATGGAATCTTCCACTCAGAGACTTCAATCATG 1141  
Db 1082 GCTCAATTTTATCCCTGAGTTTCAGAGATGGAATCTTCCACTCAGAGACTTCAATCATG 1141  
Qy 1142 CGTGTGAAGCTCACTGCTGCTTCTCATCTTGAAGATGTTAGTGTGATGCTTTGAGA 1201  
Db 1142 CGTGTGAAGCTCACTGCTGCTTCTCATCTTGAAGATGTTAGTGTGATGCTTTGAGA 1201  
Qy 1202 CATAGAGTTTGTGATATCCGCAAAAGCTCTGAAAGGTTAGGGGATTAAGAGGCTTA 1261  
Db 1202 CATAGAGTTTGTGATATCCGCAAAAGCTCTGAAAGGTTAGGGGATTAAGAGGCTTA 1261  
Qy 1262 GATAGAGAGGTGGGCTTTTGTGATGTTGAAAAATCTTAAGAAAGTTGGTACCTTTCT 1321  
Db 1262 GATAGAGAGGTGGGCTTTTGTGATGTTGAAAAATCTTAAGAAAGTTGGTACCTTTCT 1321  
Qy 1322 AGAGATTTCTGACTTGAAGATTAAGAAAAAGCCAGGCTGCAATGCTTAACACATAT 1381  
Db 1322 AGAGATTTCTGACTTGAAGATTAAGAAAAAGCCAGGCTGCAATGCTTAACACATAT 1381  
Qy 1382 AACTTGGGAACTTTAGGACAGAGGATGATTAAGTTCAAGGCTAGCCAGGCTATGCTGTA 1441  
Db 1382 AACTTGGGAACTTTAGGACAGAGGATGATTAAGTTCAAGGCTAGCCAGGCTATGCTGTA 1441  
Qy 1442 AGACTGTCTCAGCATCAAGAGCAAAATTAACATAGAGACAGAGAGGCTGAGAGTGA 1501  
Db 1442 AGACTGTCTCAGCATCAAGAGCAAAATTAACATAGAGACAGAGAGGCTGAGAGTGA 1501  
Qy 1502 GGGTCGAGACAGTGAAGGCTGATTTGTATCAAGACAGAGGAATCTATATTTGATGTAACC 1561  
Db 1502 GGGTCGAGACAGTGAAGGCTGATTTGTATCAAGACAGAGGAATCTATATTTGATGTAACC 1561  
Qy 1562 CCACATGAAAAAGCTGAGGCTGGTGAAGCATGCTTGAAGCTCAAGAGATGAGAGTAA 1621  
Db 1562 CCACATGAAAAAGCTGAGGCTGGTGAAGCATGCTTGAAGCTCAAGAGATGAGAGTAA 1621  
Db 1622 AGGCAACAAGATCCCGGGGCTTGGCTGAGTCAAGCTTAAAGCTGATGTTCCAA 1681  
Qy 1622 AGGCAACAAGATCCCGGGGCTTGGCTGAGTCAAGCTTAAAGCTGATGTTCCAA 1681  
Db 1682 GTTCCAAAGAGTCCCTGCTCAGAGAGACATGCAATGATGATGATGATGATGATGATGAT 1741  
Qy 1682 GTTCCAAAGAGTCCCTGCTCAGAGAGACATGCAATGATGATGATGATGATGATGATGAT 1741  
Db 1682 GTTCCAAAGAGTCCCTGCTCAGAGAGACATGCAATGATGATGATGATGATGATGATGAT 1741  
Qy 1742 GGTGTCTCTCTCTCTCAGAGAGACATGCAATGATGATGATGATGATGATGATGATGAT 1801  
Db 1742 GGTGTCTCTCTCTCTCAGAGAGACATGCAATGATGATGATGATGATGATGATGATGAT 1801  
Qy 1802 AC 1861  
Db 1802 AC 1861  
Qy 1862 TCTCTATTAACATGATCT 1921  
Db 1862 TCTCTATTAACATGATCT 1921  
Qy 1922 CATGCGAGAGAGTCCAGTAATTTATTCAGACACTCAGAAAGCTGGAGCAGAACCTGGA 1981  
Db 1922 CATGCGAGAGAGTCCAGTAATTTATTCAGACACTCAGAAAGCTGGAGCAGAACCTGGA 1981  
Qy 1982 GAGTTCAAGAGACAGTGTGCTCAACACTGCGACACTTCTTCAACAGAAAAAGTTTACC 2041  
Db 1982 GAGTTCAAGAGACAGTGTGCTCAACACTGCGACACTTCTTCAACAGAAAAAGTTTACC 2041  
Qy 2042 CGCAGAGAGCTGCTGCTGTAAGAGAAACCTGGGAAAGGCAAACTTGAAGTGTG 2101  
Db 2042 CGCAGAGAGCTGCTGCTGTAAGAGAAACCTGGGAAAGGCAAACTTGAAGTGTG 2101  
Qy 2102 TGTCTAAGGGGAACTGACTCAGACAACTTCTCATCTCTGAGGAACTGAGCTGTTTC 2161  
Db 2102 TGTCTAAGGGGAACTGACTCAGACAACTTCTCATCTCTGAGGAACTGAGCTGTTTC 2161  
Qy 2162 TGAACAGAGACACACCGGCTGAGTGGGACATACAGAGGAGAGCTCTTGCAGCAATTTAT 2221  
Db 2162 TGAACAGAGACACACCGGCTGAGTGGGACATACAGAGGAGAGCTCTTGCAGCAATTTAT 2221  
Qy 2222 TGTCTAAGGGGAACTGACTCAGACAACTTCTCATCTCTGAGGAACTGAGCTGTTTC 2281  
Db 2222 TGTCTAAGGGGAACTGACTCAGACAACTTCTCATCTCTGAGGAACTGAGCTGTTTC 2281  
Qy 2282 ACCTCAGTTTCACTGCTTACAGCTGCTGCGCAGTGTGATCTGTGCTGCTGCC 2341  
Db 2282 ACCTCAGTTTCACTGCTTACAGCTGCTGCGCAGTGTGATCTGTGCTGCTGCC 2341  
Qy 2342 ATCTATTAACAGATCAAAATTAATGAGACCCGAGTGAATAATTAAGTGAAGCAAAAGT 2401  
Db 2342 ATCTATTAACAGATCAAAATTAATGAGACCCGAGTGAATAATTAAGTGAAGCAAAAGT 2401  
Qy 2402 AGCTTGTCTCAAGATTTTGTGATGAGTGGGAGCACTGTATCATCAGAGACATCTGT 2461  
Db 2402 AGCTTGTCTCAAGATTTTGTGATGAGTGGGAGCACTGTATCATCAGAGACATCTGT 2461  
Qy 2462 TAGTGAAGACACCAAAACCTGTGTGATCCGTTTTTTCATGATGATGATGATGATGATGAT 2521  
Db 2462 TAGTGAAGACACCAAAACCTGTGTGATCCGTTTTTTCATGATGATGATGATGATGATGAT 2521  
Qy 2522 TGTCTTACCTGCTGTGAGAGTCTGTGCTTCTTATAGTGGTATGGAAGGAGACATC 2581  
Db 2522 TGTCTTACCTGCTGTGAGAGTCTGTGCTTCTTATAGTGGTATGGAAGGAGACATC 2581  
Qy 2582 TTAACAAATTCATGAGATTAACAGCTCATGACAGAGGAGAACTTAATCTCAATGTT 2641  
Db 2582 TTAACAAATTCATGAGATTAACAGCTCATGACAGAGGAGAACTTAATCTCAATGTT 2641  
Qy 2642 TTAAGTAAATTAACCTGATCTGCAAAAGTACTTGTGATGATGATGATGATGATGATGAT 2701  
Db 2642 TTAAGTAAATTAACCTGATCTGCAAAAGTACTTGTGATGATGATGATGATGATGATGAT 2701



QY 1262 GATAGAGAGTGGGCTTTTGTGATGTTGAAAAATCTTAAGAAGTGGTACTTTTCT 1321  
 DB 1262 GATAGAGAGTGGGCTTTTGTGATGTTGAAAAATCTTAAGAAGTGGTACTTTTCT 1321  
 QY 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGGCATATGCTTAACAGATAT 1381  
 DB 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGGCATATGCTTAACAGATAT 1381  
 QY 1382 AACTTGGGAACTTTAGGAGAGAGGTGATTAAGTTCAAGGTCAAGCCAGGGCTATGCTGTA 1441  
 DB 1382 AACTTGGGAACTTTAGGAGAGAGGTGATTAAGTTCAAGGTCAAGCCAGGGCTATGCTGTA 1441  
 QY 1442 AGACTGCTTCAMATCCAAAAGAGAAAATTAATAGACACAGACAGAGAGGCTGGAATGA 1501  
 DB 1442 AGACTGCTTCAMATCCAAAAGAGAAAATTAATAGACACAGAGAGGCTGGAATGA 1501  
 QY 1502 GGCTGGAGACAGTGGAGTGTGTATCAAGCAGAGGAATCTATATTTGATGTGAGACC 1561  
 DB 1502 GGCTGGAGACAGTGGAGTGTGTATCAAGCAGAGGAATCTATATTTGATGTGAGACC 1561  
 QY 1562 CCACATGAAAAAGCTAGAGCTGTGTAGAGCATGCTTGTAGACTCAAGAGATGAGAGGTAA 1621  
 DB 1562 CCACATGAAAAAGCTAGAGCTGTGTAGAGCATGCTTGTAGACTCAAGAGATGAGAGGTAA 1621  
 QY 1622 AGGCACAACAGATCCCGGGGGCTTGGGTGACATGCTTACCTAGGTGCTGATTCGA 1681  
 DB 1622 AGGCACAACAGATCCCGGGGGCTTGGGTGACATGCTTACCTAGGTGCTGATTCGA 1681  
 QY 1682 GTCCACAAGAGTCCCTGTCTCAMAAGTAAGATGAGCTGAGATCTGGGGCATGTCATGGG 1741  
 DB 1682 GTCCACAAGAGTCCCTGTCTCAMAAGTAAGATGAGCTGAGATCTGGGGCATGTCATGGG 1741  
 QY 1742 GGTGTCCTCTCTCTCAGAAAGAGACATGACATGWCCTGACACACACACACACACAC 1801  
 DB 1742 GGTGTCCTCTCTCTCAGAAAGAGACATGACATGWCCTGACACACACACACACACAC 1801  
 QY 1802 ACACACACACACACACACACACACACACATGAMATGAAAGTTCTCTGTGCTGCTGCTAC 1861  
 DB 1802 ACACACACACACACACACACACACACACATGAMATGAAAGTTCTCTGTGCTGCTGCTAC 1861  
 QY 1862 TCTCTTAACATATATCTCTACAGACTCTCTCTGCTTGTGTTAAGACATGAGTGGAG 1921  
 DB 1862 TCTCTTAACATATATCTCTACAGACTCTCTCTGCTTGTGTTAAGACATGAGTGGAG 1921  
 QY 1922 CATGGCAGAGCAGTCCAGTAATTAATTCAGCACTCAGAAAGCTGAGAAAGCTGGA 1981  
 DB 1922 CATGGCAGAGCAGTCCAGTAATTAATTCAGCACTCAGAAAGCTGAGAAAGCTGGA 1981  
 QY 1982 GAGTTCAAGAGCAGTGTGCCCAACCTGCCAGACTCTTCTTAACAAGAAAAAGTTACC 2041  
 DB 1982 GAGTTCAAGAGCAGTGTGCCCAACCTGCCAGACTCTTCTTAACAAGAAAAAGTTACC 2041  
 QY 2042 CGCAGAGAGCTGTGTCTTAAGAAAGAAACCTGCGAAAGGGAACCTTTGACGTGTTG 2101  
 DB 2042 CGCAGAGAGCTGTGTCTTAAGAAAGAAACCTGCGAAAGGGAACCTTTGACGTGTTG 2101  
 QY 2102 TGCTCAAGGGGAACTGACTCAGACAACTTCTCATTCTTGAAGAAACCTGAGCTGTTTC 2161  
 DB 2102 TGCTCAAGGGGAACTGACTCAGACAACTTCTCATTCTTGAAGAAACCTGAGCTGTTTC 2161  
 QY 2162 TGACAGAGAAACACCGGTGACTGGGACATACAGAGGAGAGGCTCTTGACGACATCTATA 2221  
 DB 2162 TGACAGAGAAACACCGGTGACTGGGACATACAGAGGAGAGGCTCTTGACGACATCTATA 2221  
 QY 2222 TAGTCAGCAAAATATCTTTGGGAGAGCAGTGTGCAACCAATTTGATTTCCAGCCGGTGG 2281  
 DB 2222 TAGTCAGCAAAATATCTTTGGGAGAGCAGTGTGCAACCAATTTGATTTCCAGCCGGTGG 2281  
 QY 2282 ACCTCAGTTTCATCTGGCTTACAGCTGCTGCCAGAGTCCCTTGATCTGTGCTGCC 2341  
 DB 2282 ACCTCAGTTTCATCTGGCTTACAGCTGCTGCCAGAGTCCCTTGATCTGTGCTGCC 2341

QY 2342 ATCTATAACAGATCAAAATTAATAGACCCGAGTGAAATTAATGAGCAGAAAGGT 2401  
 DB 2342 ATCTATAACAGATCAAAATTAATAGACCCGAGTGAAATTAATGAGCAGAAAGGT 2401  
 QY 2402 AGCTTTGTTCAAGATTTTTTTCATTTGGGAGACAACTGTGTACATCAGAGACATCTGT 2461  
 DB 2402 AGCTTTGTTCAAGATTTTTTTCATTTGGGAGACAACTGTGTACATCAGAGACATCTGT 2461  
 QY 2462 TAGTGAAGACACCAAAACCTGTGTACCGCTTTTTCATGTATGAATTTGTTGTTAGGT 2521  
 DB 2462 TAGTGAAGACACCAAAACCTGTGTACCGCTTTTTCATGTATGAATTTGTTGTTAGGT 2521  
 QY 2522 TGCTTCTAGCTAGCTGTGGAGGCTCGGCTTCTTAAGAGGATGAGAGGAGACATC 2581  
 DB 2522 TGCTTCTAGCTAGCTGTGGAGGCTCGGCTTCTTAAGAGGATGAGAGGAGACATC 2581  
 QY 2582 TAACAAAATCCATTAGAGATTAACAGCTCTCATGACAGAAAGGAAAATTAATCTCAATGTT 2641  
 DB 2582 TAACAAAATCCATTAGAGATTAACAGCTCTCATGACAGAAAGGAAAATTAATCTCAATGTT 2641  
 QY 2642 TTAAGTAATAAACTGTATCTGGCAAAAGTACTTTGACATTAATAAAAAAAAAAAG 2701  
 DB 2642 TTAAGTAATAAACTGTATCTGGCAAAAGTACTTTGACATTAATAAAAAAAAAAAG 2701  
 QY 2702 GGGCGCCGC 2710  
 DB 2702 GGGCGCCGC 2710  
 RESULT 5  
 AAF23458  
 ID AAF23458 standard; DNA; 2710 BP.  
 AC AAF23458;  
 XX  
 DT 21-MAR-2001 (first entry)  
 XX  
 DE Murine 200 gene.  
 XX  
 KW Cysteine protease; immune disorder; T lymphocyte; Crohn's;  
 KW arthritis; diabetes; multiple sclerosis; viral infection; bacterial;  
 KW HIV; ds.  
 XX  
 OS Mus sp.  
 XX  
 PN US6156887-A.  
 XX  
 PD 05-DEC-2000.  
 XX  
 PF 03-OCT-1997; 97US-0939729.  
 XX  
 PR 01-MAR-1996; 96US-0609583.  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Levinson DA;  
 XX  
 DR WPI; 2001-101473/11.  
 XX  
 PT Novel polypeptide exhibiting cysteine protease activity, useful for  
 PT treating and diagnosing immune disorders, especially T  
 PT lymphocyte-related disorders, e.g. Crohn's disease, multiple sclerosis,  
 PT graft versus host disease or allergies -  
 XX  
 PS Examples; Fig 17; 107bp; English.  
 CC The present invention relates to a novel cysteine protease. The  
 CC protein of the invention is useful for treating and diagnosing immune  
 CC disorders, especially T lymphocyte-related disorders. In particular,  
 CC the polypeptide is useful for treating or diagnosing T helper (TH) cell  
 CC or TH cell subpopulation-related disorders. These disorders include

CC Crohn's disease, reactive arthritis, Lyme disease, insulin-dependent  
CC diabetes, organ-specific autoimmunity, multiple sclerosis, Hashimoto's  
CC thyroiditis, Grave's disease, contact dermatitis, psoriasis, graft  
CC rejection, graft versus host disease, sarcoidosis, atopic (e.g. asthma  
CC or allergy), eosinophilia, conjunctivitis, glomerular nephritis, or  
CC helminthic (e.g. leishmaniasis), viral (e.g. HIV (human  
CC immunodeficiency virus)) or bacterial (e.g. tuberculosis or lepromatous  
CC leprosy) infections.

XX Sequence 2710 BP; 789 A; 592 C; 644 G; 678 T; 7 other;

Query Match 99.9%; Score 2706.6; DB 22; Length 2710;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 2 GTGACCCACGCGGTCCGGAATTTCCCTCCCAAGTACTATGTTTTCAGCTTACCTCA 61
Db 2 GTGACCCACGCGGTCCGGAATTTCCCTCCCAAGTACTATGTTTTCAGCTTACCTCA 61
QY 62 ACTGTCCTGCTGCTGCACTACTTGCAGAGTCATTGGAAAGTGTATTAAG 121
Db 62 ACTGTCCTGCTGCTGCACTACTTGCAGAGTCATTGGAAAGTGTATTAAG 121
QY 122 TTGAGGTGGTAAATATGCTTATCTGCTGCACTTACCTCACTCTGAGAC 181
Db 122 TTGAGGTGGTAAATATGCTTATCTGCTGCACTTACCTCACTCTGAGAC 181
QY 182 TTGTCCTATGTCGTGGGCAAGGATTTCTGCTTGGTCACAGTACATGAGTGC 241
Db 182 TTGTCCTATGTCGTGGGCAAGGATTTCTGCTTGGTCACAGTACATGAGTGC 241
QY 242 TCAGACATGATGAAGAATGTGACATATCAGAAATCCAGAGATTCAGCTTAAAGG 301
Db 242 TCAGACATGATGAAGAATGTGACATATCAGAAATCCAGAGATTCAGCTTAAAGG 301
QY 302 ATCTCAACAAAGAGATGTGTCTGATCATTAAGATGTACTGATGACATGGGA 361
Db 302 ATCTCAACAAAGAGATGTGTCTGATCATTAAGATGTACTGATGACATGGGA 361
QY 362 CCTACTGCTGCAAGTACAGTTCCTCTGCTTATGATATTAATAAATTAAGTGAAT 421
Db 362 CCTACTGCTGCAAGTACAGTTCCTCTGCTTATGATATTAATAAATTAAGTGAAT 421
QY 422 TAGACATCAAAAGCAGCAAGTCACTCCAGCTCAGACTGCCATGGGGACTACTACAG 481
Db 422 TAGACATCAAAAGCAGCAAGTCACTCCAGCTCAGACTGCCATGGGGACTACTACAG 481
QY 482 CTTCCTCAAGAACCTTAACCAAGAGAGAAATGTTTCAAGACACAGACACTGTGACCC 541
Db 482 CTTCCTCAAGAACCTTAACCAAGAGAGAAATGTTTCAAGACACAGACACTGTGACCC 541
QY 542 TCCATTAATTAACATGGAACAAAATTTCCACATGGGCTGATGATAATTAAGACTGTGAG 601
Db 542 TCCATTAATTAACATGGAACAAAATTTCCACATGGGCTGATGATAATTAAGACTGTGAG 601
QY 602 AAACGATCAAGACTGCTATCCACATTTGAGTGGAGTCTGTGGGTTGACCTGGGAC 661
Db 602 AAACGATCAAGACTGCTATCCACATTTGAGTGGAGTCTGTGGGTTGACCTGGGAC 661
QY 662 TTATCATTTGTTCTTAATCTTAATGTTATTTCTGTAAAGAAAAGATTTATGAGTT 721
Db 662 TTATCATTTGTTCTTAATCTTAATGTTATTTCTGTAAAGAAAAGATTTATGAGTT 721
QY 722 TGAAGCTTATTAACATGCGCAACTGTCTCCAGAGGGTGGCAAAATGACAGAGCTCA 781
Db 722 TGAAGCTTATTAACATGCGCAACTGTCTCCAGAGGGTGGCAAAATGACAGAGCTCA 781
QY 782 GGATTCGCTCTGAGGAAAATATCTACCATCGAGAGAAAGTATATGAAGTGAAGATT 841
Db 782 GGATTCGCTCTGAGGAAAATATCTACCATCGAGAGAAAGTATATGAAGTGAAGATT 841
QY 842 CAATGAGTACTAGTCTAGCTCAACAGCAGACAGCATCTGACCGCTCTGAGCTGCC 901
Db 842 CAATGAGTACTAGTCTAGCTCAACAGCAGACAGCATCTGACCGCTCTGAGCTGCC 901
```

```
Db 842 CAATGAGTACTAGTCTAGCTCAACAGCAGACAGCATCTGACCGCTCTGAGCTGCC 901
QY 902 ACTTTAAAGGTGGCTTCAATTTGCTTGTGATTTTCCCTTTTGTGAAAACATATGCG 961
Db 902 ACTTTAAAGGTGGCTTCAATTTGCTTGTGATTTTCCCTTTTGTGAAAACATATGCG 961
QY 962 ATATGTCATTTGGCAACCTCATTTGGAGTTTGCACACAGCCATGAGAAAAGATTGCA 1021
Db 962 ATATGTCATTTGGCAACCTCATTTGGAGTTTGCACACAGCCATGAGAAAAGATTGCA 1021
QY 1022 GTTTTCTGGGGATAATTAATCTCAAGAGGGATTCAGCTTAATCTATGCTACATTAAT 1081
Db 1022 GTTTTCTGGGGATAATTAATCTCAAGAGGGATTCAGCTTAATCTATGCTACATTAAT 1081
QY 1082 GCTTCATTTTATCCCTGAGTTTCAGAGATGGATTCCTCCACTCCAGAGACTTCAATCATG 1141
Db 1082 GCTTCATTTTATCCCTGAGTTTCAGAGATGGATTCCTCCACTCCAGAGACTTCAATCATG 1141
QY 1142 CGTGTGAAGCTCACTGCTGCTTTTCAATCAATTAAGGATGTTAGTGTGCTTTGAGA 1201
Db 1142 CGTGTGAAGCTCACTGCTGCTTTTCAATCAATTAAGGATGTTAGTGTGCTTTGAGA 1201
QY 1202 CATAGAGTTTGTGATATATCCGCAAGCTCTGAAACAGTATGAGGGAATAAGGGCTTA 1261
Db 1202 CATAGAGTTTGTGATATATCCGCAAGCTCTGAAACAGTATGAGGGAATAAGGGCTTA 1261
QY 1262 CATAGAGTTTGTGATATATCCGCAAGCTCTGAAACAGTATGAGGGAATAAGGGCTTA 1321
Db 1262 CATAGAGTTTGTGATATATCCGCAAGCTCTGAAACAGTATGAGGGAATAAGGGCTTA 1321
QY 1322 AGAGATTTCTGACCTTGAAGGATTAAGAAAAGCAGAGTGGCATATGCTTAACAGATAT 1381
Db 1322 AGAGATTTCTGACCTTGAAGGATTAAGAAAAGCAGAGTGGCATATGCTTAACAGATAT 1381
QY 1382 AACTTGGGAACCTTAGCAGAGAGGGTGAATTAAGTTCAAGGCTACAGGGCTATGCTGTA 1441
Db 1382 AACTTGGGAACCTTAGCAGAGAGGGTGAATTAAGTTCAAGGCTACAGGGCTATGCTGTA 1441
QY 1442 AGACTGTCTCAACATCCAAAGCAGAAATTAACATTAAGACACAGCAGAGGCTGAGATTA 1501
Db 1442 AGACTGTCTCAACATCCAAAGCAGAAATTAACATTAAGACACAGCAGAGGCTGAGATTA 1501
QY 1502 GACTGGAACAGTGAAGTGCATTTGTATCAAGACAGAGGATATATATTTGATGATGAC 1561
Db 1502 GACTGGAACAGTGAAGTGCATTTGTATCAAGACAGAGGATATATATTTGATGATGAC 1561
QY 1562 CCAATGGAAGAGCTAGGCTCTGTAGACATGCTTGTAGACTCAAGAGATGAGAGTAA 1621
Db 1562 CCAATGGAAGAGCTAGGCTCTGTAGACATGCTTGTAGACTCAAGAGATGAGAGTAA 1621
QY 1622 AGGCAACAAGATCCCGGGGCTTGCCTGCACTCAGCTTAAGCTTAGTGTGCTGATCCAA 1681
Db 1622 AGGCAACAAGATCCCGGGGCTTGCCTGCACTCAGCTTAAGCTTAGTGTGCTGATCCAA 1681
QY 1682 GTCCACAAGAGTCCCTGCTCAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1741
Db 1682 GTCCACAAGAGTCCCTGCTCAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1741
QY 1742 GGTGTCTCTCTCTCTCAAGAGAGCATGCAATGCCCTGACACACACACACACAC 1801
Db 1742 GGTGTCTCTCTCTCTCTCAAGAGAGCATGCAATGCCCTGACACACACACACACAC 1801
QY 1802 ACACACACACACACACACACACACACACATGATGAAGGTTCTTCTGTGCTGCTTAC 1861
Db 1802 ACACACACACACACACACACACACACACATGATGAAGGTTCTTCTGTGCTGCTTAC 1861
QY 1862 TCTCTAATCAATGATCTCTCAAGAGCTCTCTCTGCTTGTGAAGCATGAGTGGAG 1921
Db 1862 TCTCTAATCAATGATCTCTCAAGAGCTCTCTCTGCTTGTGAAGCATGAGTGGAG 1921
QY 1922 CATGGCAGAGAGTCCAGTAAATTTATTCACAGACTCAAGAGGCTGAGCAGAGGCTGGA 1981
Db 1922 CATGGCAGAGAGTCCAGTAAATTTATTCACAGACTCAAGAGGCTGAGCAGAGGCTGGA 1981
```



QY 302 ATCTCAACAAAGAGATGTGTCTGTGATCATTAAGAATGTGACTGTGATGACCATGAGGA 361  
| | | | |  
Db 302 ATCTCAACAAAGAGATGTGTCTGTGATCATTAAGAATGTGACTGTGATGACCATGAGGA 361  
QY 362 CTTACTGCTGCAGGATACAGTTCCTGGTCTTATGATGATTAATAAATTAGAACTGAAT 421  
| | | | |  
Db 362 CTTACTGCTGCAGGATACAGTTCCTGGTCTTATGATGATTAATAAATTAGAACTGAAT 421  
QY 422 TAGACATCAAAAGCAGCAGGATCACTCCAGCTCAGACTGGCCCATGAGGAGCTTACTTAAG 481  
| | | | |  
Db 422 TAGACATCAAAAGCAGCAGGATCACTCCAGCTCAGACTGGCCCATGAGGAGCTTACTTAAG 481  
QY 482 CTTCTCCAAAGAACCTTAACACGAGAGAAATGTTTACAGACACAGACACTGGTGAACC 541  
| | | | |  
Db 482 CTTCTCCAAAGAACCTTAACACGAGAGAAATGTTTACAGACACAGACACTGGTGAACC 541  
QY 542 TCCATTAATTAACATGGAACAAAATTTCCACATGAGGCTGATGAATTAAGAATCTGAG 601  
| | | | |  
Db 542 TCCATTAATTAACATGGAACAAAATTTCCACATGAGGCTGATGAATTAAGAATCTGAG 601  
QY 602 AAACGATCAAGACCTGCTATCCATTTGAGAGTGGAGTCTGCTGGTGAACCTTGAC 661  
| | | | |  
Db 602 AAACGATCAAGACCTGCTATCCATTTGAGAGTGGAGTCTGCTGGTGAACCTTGAC 661  
QY 662 TTATCATTTGCTCTTAATCTTAAATGGTATTCCTGTGAAGAAAGAAAGTATGAGT 721  
| | | | |  
Db 662 TTATCATTTGCTCTTAATCTTAAATGGTATTCCTGTGAAGAAAGAAAGTATGAGT 721  
QY 722 TGAGCTTATTAACACTGAGCAACTTGGCTCCAGAGAGGTTGGCAATGACAGAGCACTCA 781  
| | | | |  
Db 722 TGAGCTTATTAACACTGAGCAACTTGGCTCCAGAGAGGTTGGCAATGACAGAGCACTCA 781  
QY 782 GGAATGCTGAGAGAGAAATATCTACACATGAGAGAGAACTATATGAAAGTGAAT 841  
| | | | |  
Db 782 GGAATGCTGAGAGAGAAATATCTACACATGAGAGAGAACTATATGAAAGTGAAT 841  
QY 842 CAATGAGTACTACTGCTACGTCACAGCAGCAGCAGCATCTGACCGCTCTGACCTGCC 901  
| | | | |  
Db 842 CAATGAGTACTACTGCTACGTCACAGCAGCAGCAGCATCTGACCGCTCTGACCTGCC 901  
QY 902 ACTTTTAAAGGCTGCTTCAATTTCTGACTTTGTATTTCCCTTTTATGAGAAATATG 961  
| | | | |  
Db 902 ACTTTTAAAGGCTGCTTCAATTTCTGACTTTGTATTTCCCTTTTATGAGAAATATG 961  
QY 962 ATATGTCACCTTGGCAACTTATGAGGTTCTGACACACAGCCTGAGAGAAAGATTTCA 1021  
| | | | |  
Db 962 ATATGTCACCTTGGCAACTTATGAGGTTCTGACACACAGCCTGAGAGAAAGATTTCA 1021  
QY 1022 GTTTCTGGGAGTAATTAATCAACAAAGGGAATTCAGCTGTAACTAGCTACATTAAG 1081  
| | | | |  
Db 1022 GTTTCTGGGAGTAATTAATCAACAAAGGGAATTCAGCTGTAACTAGCTACATTAAG 1081  
QY 1082 GCTCATATTTTATCCCTGAGTTTCAAGGATCGGATCTCCCATCCAGAGACTTCAATCAG 1141  
| | | | |  
Db 1082 GCTCATATTTTATCCCTGAGTTTCAAGGATCGGATCTCCCATCCAGAGACTTCAATCAG 1141  
QY 1142 CGTGTGAAAGTCACTCGCTGCTTATATATAGAAATGATGATGATGATGATGATGATG 1201  
| | | | |  
Db 1142 CGTGTGAAAGTCACTCGCTGCTTATATATAGAAATGATGATGATGATGATGATGATG 1201  
QY 1202 CATAGAGTTTGTGATATCCGCAAGGCTCTGAAACAGGTAGGAGGAAATTAAGGAGCTAA 1261  
| | | | |  
Db 1202 CATAGAGTTTGTGATATCCGCAAGGCTCTGAAACAGGTAGGAGGAAATTAAGGAGCTAA 1261  
QY 1262 GATAGAGAGTGCCTGCTTGTGATGTTGAAAAATCTTAAGAAAGTGTGATGCTTTCT 1321  
| | | | |  
Db 1262 GATAGAGAGTGCCTGCTTGTGATGTTGAAAAATCTTAAGAAAGTGTGATGCTTTCT 1321  
QY 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAGCAGAGTGCATATGCTTAACAGATAT 1381  
| | | | |  
Db 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAGCAGAGTGCATATGCTTAACAGATAT 1381

QY 1382 AACTTGGAGACCTTAGGAGGAGGTGATAGTTCAAGGTCAAGCCAGGCTATGCTGATA 1441  
| | | | |  
Db 1382 AACTTGGAGACCTTAGGAGGAGGTGATAGTTCAAGGTCAAGCCAGGCTATGCTGATA 1441  
QY 1442 AGACTGTCTCAACATCCAAAGACGAAATTAACATAGACACAGAGGCTGAGATGA 1501  
| | | | |  
Db 1442 AGACTGTCTCAACATCCAAAGACGAAATTAACATAGACACAGAGGCTGAGATGA 1501  
QY 1502 GGGCTGGAGCAGTGGAGTGCATTTGTGTACAGACAGAGAAATCTATATGATTCGAGACC 1561  
| | | | |  
Db 1502 GGGCTGGAGCAGTGGAGTGCATTTGTGTACAGACAGAGAAATCTATATGATTCGAGACC 1561  
QY 1562 CCACATGAAAAAGCTAGGCTGTAGAGCATGCTGTGTAGACTCAAGAGATGAGAGGTAA 1621  
| | | | |  
Db 1562 CCACATGAAAAAGCTAGGCTGTAGAGCATGCTGTGTAGACTCAAGAGATGAGAGGTAA 1621  
QY 1622 AGGCAACAACAGATCCCGGGGCTTGCGTGCAGTCACTTAAGCTTGAAGTTCCAA 1681  
| | | | |  
Db 1622 AGGCAACAACAGATCCCGGGGCTTGCGTGCAGTCACTTAAGCTTGAAGTTCCAA 1681  
QY 1682 GTCCACAAAGATCCCTGCTCAGAAAGTATGATGATGATGATGATGATGATGATGATG 1741  
| | | | |  
Db 1682 GTCCACAAAGATCCCTGCTCAGAAAGTATGATGATGATGATGATGATGATGATGATG 1741  
QY 1742 GATTGTCTCTCTCTCAGAAAGACATGACATGACCTGCAACACACACACACACAC 1801  
| | | | |  
Db 1742 GATTGTCTCTCTCTCAGAAAGACATGACATGACCTGCAACACACACACACACAC 1801  
QY 1802 AACACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATG 1861  
| | | | |  
Db 1802 AACACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATG 1861  
QY 1862 TCTCTATTAACATGATCTCTCAGAGACCTCTCTGCTCTGTTAAAGCATGAGGAG 1921  
| | | | |  
Db 1862 TCTCTATTAACATGATCTCTCAGAGACCTCTCTGCTCTGTTAAAGCATGAGGAG 1921  
QY 1922 CATGACAGAGACATGCTCAATTAATTCACACACTCAGAAAGCTGAGACAGAGGTTGA 1981  
| | | | |  
Db 1922 CATGACAGAGACATGCTCAATTAATTCACACACTCAGAAAGCTGAGACAGAGGTTGA 1981  
QY 1982 GAGTTACAGAGACATGCTGCTCAACACATGCTCTTCTTACACAGAAAAAGTTACC 2041  
| | | | |  
Db 1982 GAGTTACAGAGACATGCTGCTCAACACATGCTCTTCTTACACAGAAAAAGTTACC 2041  
QY 2042 CGCAAGCAGCCTGCTGTGTGTAAGAAAGAACCTGCGCAAGAAAGCAATTTGATGTTG 2101  
| | | | |  
Db 2042 CGCAAGCAGCCTGCTGTGTGTAAGAAAGAACCTGCGCAAGAAAGCAATTTGATGTTG 2101  
QY 2102 TGCTCAAGGGGACATGACTCAGACCACTTCTCAATCTCTGAGAGGAACTGAGAGCTGTT 2161  
| | | | |  
Db 2102 TGCTCAAGGGGACATGACTCAGACCACTTCTCAATCTCTGAGAGGAACTGAGAGCTGTT 2161  
QY 2162 TGACAGAGAAACAACCGGTGACTGGGACATAGAGGAGAGGCTTTGACAGCAATCTATA 2221  
| | | | |  
Db 2162 TGACAGAGAAACAACCGGTGACTGGGACATAGAGGAGAGGCTTTGACAGCAATCTATA 2221  
QY 2222 TAGTCAGAAAAATATCTTTGGAGAGACATGCTCAACAAATGATTTTCAAGCCGGTGG 2281  
| | | | |  
Db 2222 TAGTCAGAAAAATATCTTTGGAGAGACATGCTCAACAAATGATTTTCAAGCCGGTGG 2281  
QY 2282 ACCTCAGTTTATCTGAGCTTACAGCTGCTGCCAGTGCCTTGAATCTGTGCTGCC 2341  
| | | | |  
Db 2282 ACCTCAGTTTATCTGAGCTTACAGCTGCTGCCAGTGCCTTGAATCTGTGCTGCC 2341  
QY 2342 ATCTATAACAGAAATCAATTAATAGACCCGAGTGAAATATTAAGTACAGAAAGT 2401  
| | | | |  
Db 2342 ATCTATAACAGAAATCAATTAATAGACCCGAGTGAAATATTAAGTACAGAAAGT 2401  
QY 2402 AGCTTTGTTCAAGATTTTGTGATTTGGAGAGCACTGTATCACTCAGAGACATCTGT 2461  
| | | | |  
Db 2402 AGCTTTGTTCAAGATTTTGTGATTTGGAGAGCACTGTATCACTCAGAGACATCTGT 2461  
QY 2462 TAGTGAGACACCAAAACCTGTGTATCGTTTTTCATGTATGAATTTTGTGTTAGGT 2521  
| | | | |







```
|||||
Db      542 TCCATATATACATGGAACAAAAATTTCCACATGGCTGATGAATTAAGACTCTGGAG 601
Qy      602 AAAAGATCGAAGCTGATCCACATTTGAGTGGAGTCTCTGGTGGACCCCTGGAC 661
Db      602 AAAAGATCGAAGCTGATCCACATTTGAGTGGAGTCTCTGGTGGACCCCTGGAC 661
Qy      662 TTATCATTTGATGCTCTTAATCCTTAAATGATATTCCTGTAAAGAAAAAGAAAGTATGAGTT 721
Db      662 TTATCATTTGATGCTCTTAATCCTTAAATGATATTCCTGTAAAGAAAAAGAAAGTATGAGTT 721
Qy      722 TGAGCTTTATTAACCTGGCCAACTTGCTCCAGAGAGGTTGGCAATGACAGAGCATCA 781
Db      722 TGAGCTTTATTAACCTGGCCAACTTGCTCCAGAGAGGTTGGCAATGACAGAGCATCA 781
Qy      782 GGATTCGCTCTAGAGAAAAATATCTACACATCGAGAGAAAGTATTTGAAGTGGAGAAAT 841
Db      782 GGATTCGCTCTAGAGAAAAATATCTACACATCGAGAGAAAGTATTTGAAGTGGAGAAAT 841
Qy      842 CAATGAGTACTACTGCTACGTCACACAGCAGAGCATCTGACCGGCTCTGAGCTGCC 901
Db      842 CAATGAGTACTACTGCTACGTCACACAGCAGAGCATCTGACCGGCTCTGAGCTGCC 901
Qy      902 ACTTTAAAGGCTCGCTTCATTTCTGACTTTGTAATTTCCCTTTXTGAAAACTATGTG 961
Db      902 ACTTTAAAGGCTCGCTTCATTTCTGACTTTGTAATTTCCCTTTXTGAAAACTATGTG 961
Qy      962 AATATGCACTTGGCAACCTCATTTGAGAGTTCGACACAGCCCATGAGAAAAAGTTCGA 1021
Db      962 AATATGCACTTGGCAACCTCATTTGAGAGTTCGACACAGCCCATGAGAAAAAGTTCGA 1021
Qy      1022 GTTTTCTGGGGATTAATTACTCAACAAGGGATTCGACTGAATCTCATGTTAGTAAT 1081
Db      1022 GTTTTCTGGGGATTAATTACTCAACAAGGGATTCGACTGAATCTCATGTTAGTAAT 1081
Qy      1082 GCTCATTTTATCCCTGAGTTTCAGGGATCGGATCTCCCATCCAGAGACTTCAATCATG 1141
Db      1082 GCTCATTTTATCCCTGAGTTTCAGGGATCGGATCTCCCATCCAGAGACTTCAATCATG 1141
Qy      1142 CGTGTGAAGCTCACTCGTCTTTCAATCATTAAGAAATGTTAGTGTCTTTGAGA 1201
Db      1142 CGTGTGAAGCTCACTCGTCTTTCAATCATTAAGAAATGTTAGTGTCTTTGAGA 1201
Qy      1202 CATGAGGTTTGTGATATCCGCAAGGCTCGTGAACAGGTGAGGGAAATTAAGGCTGA 1261
Db      1202 CATGAGGTTTGTGATATCCGCAAGGCTCGTGAACAGGTGAGGGAAATTAAGGCTGA 1261
Qy      1262 GATAGGAAGTCCGCTTTTGTGATGTTGAAAAATCTTAAAGAAATGTTAGTGTCT 1321
Db      1262 GATAGGAAGTCCGCTTTTGTGATGTTGAAAAATCTTAAAGAAATGTTAGTGTCT 1321
Qy      1322 AAGAAATTTCTGACCTTTGAAAGATTAAGAAAAAGCCAGGTGGCATATGCTTAAACAGATAT 1381
Db      1322 AAGAAATTTCTGACCTTTGAAAGATTAAGAAAAAGCCAGGTGGCATATGCTTAAACAGATAT 1381
Qy      1382 AACTTGGGAACCTTAGCAGAGAGGATGAAGTTCAAGGTCAGCCAGGCTATGCTGTA 1441
Db      1382 AACTTGGGAACCTTAGCAGAGAGGATGAAGTTCAAGGTCAGCCAGGCTATGCTGTA 1441
Qy      1442 AAGTGTCTCAACATCAAAAGAGAAATTAACATAGAGACAGAGAGGCTGAGATGA 1501
Db      1442 AAGTGTCTCAACATCAAAAGAGAAATTAACATAGAGACAGAGAGGCTGAGATGA 1501
Qy      1502 GGCTCGGACAGTGAAGTGTGTGTACAAGCAAGAAATCTATATTTGATGCTGAGACC 1561
Db      1502 GGCTCGGACAGTGAAGTGTGTGTACAAGCAAGAAATCTATATTTGATGCTGAGACC 1561
Qy      1562 CCACATGAAAAAGCTGAGGCTGTGAGAGATGCTTGAAGCTCAAGAAATGGAAGGTA 1621
Db      1562 CCACATGAAAAAGCTGAGGCTGTGAGAGATGCTTGAAGCTCAAGAAATGGAAGGTA 1621
Qy      1622 AAGCAACAAGATCCCGGGGCTTGCTGCAAGTCAAGCTTAGCTGAGTTCCAA 1681
Db      1622 AAGCAACAAGATCCCGGGGCTTGCTGCAAGTCAAGCTTAGCTGAGTTCCAA 1681

|||||
Db      1682 AAGCAACAAGATCCCGGGGCTTGCTGCAAGTCAAGCTTAGCTGAGTTCCAA 1681
Qy      1682 GTCCACAAGAGTCCCTGTCTGAMAGTAGATGAGTCTGAGTCTGGCCATGTCATGGG 1741
Db      1682 GTCCACAAGAGTCCCTGTCTGAMAGTAGATGAGTCTGAGTCTGGCCATGTCATGGG 1741
Qy      1742 GGTGTCTCTCTCTCAGAAAGAGATGACATGATCCCTGACACACACACACACAC 1801
Db      1742 GGTGTCTCTCTCTCAGAAAGAGATGACATGATCCCTGACACACACACACACAC 1801
Qy      1802 ACACACACACACACACACACACACACATGATGAATGAATGTTCTCTGTGCTCTAC 1861
Db      1802 ACACACACACACACACACACACACACATGATGAATGAATGTTCTCTGTGCTCTAC 1861
Qy      1862 TCTCTATTAACATGATCTCTCAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921
Db      1862 TCTCTATTAACATGATCTCTCAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921
Qy      1922 CATGSCAGAGAGTCCAGTAATTTATTCACAGACTCAGAAAGGCTGAGCAGAGGTGA 1981
Db      1922 CATGSCAGAGAGTCCAGTAATTTATTCACAGACTCAGAAAGGCTGAGCAGAGGTGA 1981
Qy      1982 GAGTTCAGGAGCACTGTGCCCAACACTGCCAGACTCTTTTACACAGAAAAAGTTACC 2041
Db      1982 GAGTTCAGGAGCACTGTGCCCAACACTGCCAGACTCTTTTACACAGAAAAAGTTACC 2041
Qy      2042 CGCAAGCAGCCTGCTGTCTGTAAGAAAGAAACCTGCGAAAGCAACTTGTGCTGTGG 2101
Db      2042 CGCAAGCAGCCTGCTGTCTGTAAGAAAGAAACCTGCGAAAGCAACTTGTGCTGTGG 2101
Qy      2102 TGCTCAAGGGGAACTGACTCAGACAACTTCTCATTTCTGAGAGAACTGAGCTGTTTC 2161
Db      2102 TGCTCAAGGGGAACTGACTCAGACAACTTCTCATTTCTGAGAGAACTGAGCTGTTTC 2161
Qy      2162 TGACAGAAAGAACACCGGTGACTGGGACATAGAAAGGACAGAGCTTTCAGCAATCTATA 2221
Db      2162 TGACAGAAAGAACACCGGTGACTGGGACATAGAAAGGACAGAGCTTTCAGCAATCTATA 2221
Qy      2222 TAGTACAGAAAAATATCTTTTGGAGAGACAGTGTGTACCAATTTGATTTCCAAAGCCGGTGG 2281
Db      2222 TAGTACAGAAAAATATCTTTTGGAGAGACAGTGTGTACCAATTTGATTTCCAAAGCCGGTGG 2281
Qy      2282 ACCTCAGTTTCACTGGCTTACAGCTGCTGCCAGTCCCTGATCTGTGCTGCC 2341
Db      2282 ACCTCAGTTTCACTGGCTTACAGCTGCTGCCAGTCCCTGATCTGTGCTGCC 2341
Qy      2342 ATCTATTAACAGAAATCAAAATTAATAGACCCGAGTGAAATATTAAGTACAGAAAGGT 2401
Db      2342 ATCTATTAACAGAAATCAAAATTAATAGACCCGAGTGAAATATTAAGTACAGAAAGGT 2401
Qy      2402 AGCTTTGTTCAAAAGTTTTTTTGGCATTTGGGAGCAACTGTGTACATCAGAGGACATCTGT 2461
Db      2402 AGCTTTGTTCAAAAGTTTTTTTGGCATTTGGGAGCAACTGTGTACATCAGAGGACATCTGT 2461
Qy      2462 TAGTGAAGCAACCAAAACCTGTGTAACGTTTTTTCATGTATGAATTTTGTGTTAGGT 2521
Db      2462 TAGTGAAGCAACCAAAACCTGTGTAACGTTTTTTCATGTATGAATTTTGTGTTAGGT 2521
Qy      2522 TGCTTTCTAGCTAGCTGTGAGAGTCTGTGCTTTCTTAAGTGGTATGGAAGGAGACATC 2581
Db      2522 TGCTTTCTAGCTAGCTGTGAGAGTCTGTGCTTTCTTAAGTGGTATGGAAGGAGACATC 2581
Qy      2582 TTAACAAATCCCTTTGAGATTAACAGCTCTCATGCAAGAAAGGAAATCTAATCTCAATGTT 2641
Db      2582 TTAACAAATCCCTTTGAGATTAACAGCTCTCATGCAAGAAAGGAAATCTAATCTCAATGTT 2641
Qy      2642 TTAAGTATATAAACTGACTGCGCAAGATCTTGAAGCATTAATAAAAAAAAAAAAAAAG 2701
Db      2642 TTAAGTATATAAACTGACTGCGCAAGATCTTGAAGCATTAATAAAAAAAAAAAAAAAG 2701
Qy      2702 GGGGGCGC 2710
Db      2702 GGGGGCGC 2710
```

```
RESULT 8
AAT38265
ID AAT38265 standard; cDNA; 2712 BP.
XX
AC AAT38265;
XX
DT 29-DEC-1996 (first entry)
XX
DE 200 gene differentially expressed in T helper cells.
XX
KM T helper cell; TH cell; T-cell; T-lymphocyte; 200 gene;
KM differential expression; immune disorder; multiple sclerosis;
KM asthma; lepromatous leprosy; diagnosis; therapy; receptor; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT misc_difference 1 /*tag= a
FT /note= "unidentified base"
FT CDS 40..885
FT /*tag= b
FT sig_peptide 40..102
FT /*tag= c
FT mat_peptide 103..882
FT /*tag= d
FT misc_difference 947
FT /*tag= e
FT /note= "base 946 is given as k in the
FT specification"
FT misc_difference 1323
FT /*tag= f
FT /note= "the identity of base 1322 is unclear in
FT Fig 17"
FT misc_difference 1455
FT /*tag= g
FT /note= "base 1454 is given as m in the
FT specification"
FT misc_difference 1706
FT /*tag= h
FT /note= "base 1705 is given as m in the
FT specification"
FT misc_difference 1717
FT /*tag= i
FT /note= "base 1716 is given as r in the
FT specification"
FT misc_difference 1779
FT /*tag= j
FT /note= "base 1778 is given as w in the
FT specification"
FT misc_difference 1836
FT /*tag= k
FT /note= "base 1835 is given as w in the
FT specification"
XX
XX WO9627603-A1.
XX
XX 12-SEP-1996.
XX
XX 01-MAR-1996; 96MO-US02798.
XX
XX 07-JUN-1995; 95US-0487748.
XX
XX 03-MAR-1995; 95US-0398633.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Levinson DA;
XX
XX WPI; 1996-433404/43.
XX
XX P-PSDB; AAM01047.
XX
```

```
PT Genes and their products differentially expressed in T helper cells
PT - useful in diagnosis and treatment of immune disorders, e.g.
PT multiple sclerosis, asthma, lepromatous leprosy, etc.
XX
PS Claim 1; Fig 17; 218pp; English.
XX
CC Novel murine genes 10, 57, 105, 106, 161, 200 and 54 (AAT38260-66)
CC were identified in transgenic T-cell paradigm and T-cell clone
CC paradigm searches for genes which are differentially expressed
CC within and/or among helper T (TH) cell subpopulations. The 200
CC gene (AAT38265) is expressed at levels many-fold higher in TH1
CC cell subpopulations than in TH2 cell subpopulations. The 200 gene
CC and its product (AAM01047), a receptor of the Ig superfamily class,
CC can be used therapeutically to ameliorate immune disorders and in
CC the diagnosis of such disorders. The gene has also been used to
CC isolate the human 200 gene (AAT38267), to produce soluble 200 gene-Ig
CC fusions and to breed transgenic animals useful as models of TH cell
CC subpopulation-related disorders.
XX
SQ Sequence 2712 BP; 789 A; 592 C; 644 G; 679 T; 8 other;
XX
Query Match 99.1%; Score 2684.6; DB 17; Length 2712;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2709; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
XX
QY 2 GTGACCCACGGGTCCGATTTCCCTCCCAAGTACTATGTTTAAAGTCTTACCTCA 61
DB 2 GTGACCCACGGGTCCGATTTCCCTCCCAAGTACTATGTTTAAAGTCTTACCTCA 61
QY 62 ACTGTGCTGCTGCTGCTGCAACTACTTGCAGAGTCATTGGAAGTGTATTAAG 121
DB 62 ACTGTGCTGCTGCTGCTGCAACTACTTGCAGAGTCATTGGAAGTGTATTAAG 121
QY 122 TTGAGGTGTAAATAATGCTTATCTGCCCTGCAATTCATCTACCTATCGGACAC 181
DB 122 TTGAGGTGTAAATAATGCTTATCTGCCCTGCAATTCATCTACCTATCGGACAC 181
QY 182 TTGTGCTTATGTCCTGCGGCAAGGATTTCTGTCCTTGTCACAGTATCAATGAGTTC 241
DB 182 TTGTGCTTATGTCCTGCGGCAAGGATTTCTGTCCTTGTCACAGTATCAATGAGTTC 241
QY 242 TCAGACTGATGAAGAATGTGACATATCAGAAATCCAGAGATCCAGCTAAAGGGG 301
DB 242 TCAGACTGATGAAGAATGTGACATATCAGAAATCCAGAGATCCAGCTAAAGGGG 301
QY 302 ATCTCAACAAGAGATGTGTCTGTGATCAATAAGATGTGACTGTGATGACATGGGA 361
DB 302 ATCTCAACAAGAGATGTGTCTGTGATCAATAAGATGTGACTGTGATGACATGGGA 361
QY 362 CCTACTGCTGCAAGATACAGTTCCTGCTTATGATTAATAAATTAGAACTGAAAT 421
DB 362 CCTACTGCTGCAAGATACAGTTCCTGCTTATGATTAATAAATTAGAACTGAAAT 421
QY 422 TAGACATCAAGACAGCAAGTCACTCCAGCTCAGACTCCCATGGGGACTTACTACAG 481
DB 422 TAGACATCAAGACAGCAAGTCACTCCAGCTCAGACTCCCATGGGGACTTACTACAG 481
QY 482 CTTCTCCAGAAACCTTAACACAGAGAGAAATGTTCAAGACACAGACACTGGTACCC 541
DB 482 CTTCTCCAGAAACCTTAACACAGAGAGAAATGTTCAAGACACAGACACTGGTACCC 541
QY 542 TCCATTAATAACAATGAACAAAAATTTCCACATGGGCTGATGAATTAAGACTGTGAG 601
DB 542 TCCATTAATAACAATGAACAAAAATTTCCACATGGGCTGATGAATTAAGACTGTGAG 601
QY 602 AAACGATCAGAACTGCTATCCATTGAGTGGAGTCTCTGCTGGGTGACCCCTGGAC 661
DB 602 AAACGATCAGAACTGCTATCCATTGAGTGGAGTCTCTGCTGGGTGACCCCTGGAC 661
QY 662 TTATCATTTGGTGTCTTAATCCTTAATGATATTCCTTAAGAAAAAGATTATGAGTT 721
DB 662 TTATCATTTGGTGTCTTAATCCTTAATGATATTCCTTAAGAAAAAGATTATGAGTT 721
```

QY 722 TGAGCTTATTAACATGAGCACTTGGCTCCAGAGAGGTTGGCAAAATGACAGAGAGTCA 781  
 Db 722 TGAGCTTATTAACATGAGCACTTGGCTCCAGAGAGGTTGGCAAAATGAGAGAGTCA 781  
 QY 782 GGAATTCCTCTGAGAGAAAATATCTACACCATGAGAGAAAGTATATGAAATGAGAAAT 841  
 Db 782 GGAATTCCTCTGAGAGAAAATATCTACACCATGAGAGAAAGTATATGAAATGAGAAAT 841  
 QY 842 CAAAAGAGATCTACAGCTACGTCACAGAGAGAGTCTGACCGCTCTGAGAGTGC 901  
 Db 842 CAAAAGAGATCTACAGCTACGTCACAGAGAGAGTCTGACCGCTCTGAGAGTGC 901  
 QY 902 ACTTTTAAAGGCTCGCTCATTTCTGACTTTGTTATTTCCCTTTTGGAAAACTATGTC 961  
 Db 902 ACTTTTAAAGGCTCGCTCATTTCTGACTTTGTTATTTCCCTTTTGGAAAACTATGTC 961  
 QY 962 ATATGCTCACTTGGCAACCTCATTTGAGGTTCTGACCAAGCACTGAGAAAAAGTTCCA 1021  
 Db 962 ATATGCTCACTTGGCAACCTCATTTGAGGTTCTGACCAAGCACTGAGAAAAAGTTCCA 1021  
 QY 1022 GTTTTCTGGGGAATTAATTAACAGAGGAGATTCGATTAATCTGATTAATTAATTAAT 1081  
 Db 1022 GTTTTCTGGGGAATTAATTAACAGAGGAGATTCGATTAATCTGATTAATTAATTAAT 1081  
 QY 1082 GCTCAATTTATCCCTGAGTTTCAGGAGATCGATCTCCCATCCAGAGACTTCAATCATG 1141  
 Db 1082 GCTCAATTTATCCCTGAGTTTCAGGAGATCGATCTCCCATCCAGAGACTTCAATCATG 1141  
 QY 1142 CGTGTGAAGCTCACTCGTCTTTCATACATTAAGAAATGTTAGTGTCTTTGAGA 1201  
 Db 1142 CGTGTGAAGCTCACTCGTCTTTCATACATTAAGAAATGTTAGTGTCTTTGAGA 1201  
 QY 1202 CATAGAGTTTGTGTATATCCGCAAAAGTCTTGAAGAGAGTTGATGATGATGATGATG 1261  
 Db 1202 CATAGAGTTTGTGTATATCCGCAAAAGTCTTGAAGAGAGTTGATGATGATGATGATG 1261  
 QY 1262 GATAGAGAGTGCCTGTTCTTTGTTGATGTTGAAATCTTAAAGAGTTGATGATGATG 1321  
 Db 1262 GATAGAGAGTGCCTGTTCTTTGTTGATGTTGAAATCTTAAAGAGTTGATGATGATG 1321  
 QY 1321 T-AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCAGGTGATATCTTAACAGAT 1379  
 Db 1322 TAAAGATTTCTGACCTTGAAGATTAAGAAAAAGCAGGTGATATCTTAACAGAT 1381  
 QY 1380 ATTAATCTTGGGAATCTTAAGCAGAGAGGATTAATGATTAAGTCAAGGAGGATTAATCT 1439  
 Db 1382 ATTAATCTTGGGAATCTTAAGCAGAGAGGATTAATGATTAAGTCAAGGAGGATTAATCT 1441  
 QY 1440 TAAAGCTGTCTCAGCATCCAAAGACGAAAAATAACTAAGACAGAGAGGCTGAGAT 1499  
 Db 1442 TAAAGCTGTCTCAGCATCCAAAGACGAAAAATAACTAAGACAGAGAGGCTGAGAT 1501  
 QY 1500 GAGGCTCGACAGTGAAGTGTATTAAGCAAGCAGAGAAATCTATATTGATCTGATGA 1559  
 Db 1502 GAGGCTCGACAGTGAAGTGTATTAAGCAAGCAGAGAAATCTATATTGATCTGATGA 1561  
 QY 1560 CCCCAGATGAAAAAGTGAAGGCTGTGATGAGATCTTGTATGATCAAGAGATGAGAGAT 1619  
 Db 1562 CCCCAGATGAAAAAGTGAAGGCTGTGATGAGATCTTGTATGATCAAGAGATGAGAGAT 1621  
 QY 1620 AAAAGCAGACAGATCCCGGGGCTTGGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1679  
 Db 1622 AAAAGCAGACAGATCCCGGGGCTTGGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1681  
 QY 1680 AAGTCAACAAGATCCCTGTCTCAGAGTAAATGATGATGATGATGATGATGATGATG 1739  
 Db 1682 AAGTCAACAAGATCCCTGTCTCAGAGTAAATGATGATGATGATGATGATGATGATG 1741  
 QY 1740 GGGGTGTCTCTCTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1799  
 Db 1742 GGGGTGTCTCTCTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1801  
 QY 1800 ACACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATG 1859

Db 1802 ACACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATG 1861  
 QY 1860 CCTCTCTTAATCATATATCTCTACAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1919  
 Db 1862 CCTCTCTTAATCATATATCTCTACAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
 QY 1920 AGCATGAGAGAGTCAAGTAAATTAATTTTCCAGACTGAGAGAGGCTGAGAGAGAGT 1979  
 Db 1922 AGCATGAGAGAGTCAAGTAAATTAATTTTCCAGACTGAGAGAGGCTGAGAGAGAGT 1981  
 QY 1980 GAGAGTCAAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCT 2039  
 Db 1982 GAGAGTCAAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCT 2041  
 QY 2040 CCGGAGAGAGTCT 2099  
 Db 2042 CCGGAGAGAGTCT 2101  
 QY 2100 TGTGCTCAAGGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 2159  
 Db 2102 TGTGCTCAAGGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 2161  
 QY 2160 TCTGACAGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 2219  
 Db 2162 TCTGACAGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 2221  
 QY 2220 TATAGTCAAGAGAGATTTCTTTGGAGAGAGTCTGAGAGAGTCTGAGAGAGTCT 2279  
 Db 2222 TATAGTCAAGAGAGATTTCTTTGGAGAGAGTCTGAGAGAGTCTGAGAGAGTCT 2281  
 QY 2280 GGAAGCTCAAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCT 2339  
 Db 2282 GGAAGCTCAAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCT 2341  
 QY 2340 CCATCTATTAACAGAGATTAATTAATTAAGTAAAGTAAATTAATTAAGTAAAGTAA 2399  
 Db 2342 CCATCTATTAACAGAGATTAATTAATTAAGTAAAGTAAATTAATTAAGTAAAGTAA 2401  
 QY 2400 GTAGCTTTGTTAAAGATTTTCTTTGGAGAGAGTCTGAGAGAGTCTGAGAGAGT 2459  
 Db 2402 GTAGCTTTGTTAAAGATTTTCTTTGGAGAGAGTCTGAGAGAGTCTGAGAGAGT 2461  
 QY 2460 GTTAGAGAGAGACCAAAACCTGTGTAACCTTTTCTCATGATGATGATGATGATG 2519  
 Db 2462 GTTAGAGAGAGACCAAAACCTGTGTAACCTTTTCTCATGATGATGATGATGATG 2521  
 QY 2520 GTTGTCTTCTAGCTGAGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 2579  
 Db 2522 GTTGTCTTCTAGCTGAGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 2581  
 QY 2580 TCTTAACAAATCATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 2639  
 Db 2582 TCTTAACAAATCATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 2641  
 QY 2640 TTTTAAAGTAAATCACTGATGAGAGAGTCTTGAAGATTAAGATTAAGATTAAGAT 2699  
 Db 2642 TTTTAAAGTAAATCACTGATGAGAGAGTCTTGAAGATTAAGATTAAGATTAAGAT 2701  
 QY 2700 AGGGGGGGGGC 2710  
 Db 2702 AGGGGGGGGGC 2712  
 RESULT 9  
 ABZ68328  
 ID ABZ68328 standard; DNA; 2725 BP.  
 XX  
 AC ABZ68328;  
 XX  
 DT 22-APR-2003 (first entry)  
 XX  
 DB Nucleotide sequence of murine TIM-3 BALB/c allele.

XX T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;  
KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;  
KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;  
KW allergic T cell response; autoimmune disease; gene; ss.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT CDS /tag= a  
FT /product= "TIM-3"  
PN MO2003002722-A2.  
PD 09-JUN-2003.  
PF 01-JUL-2002; 2002MO-US20890.  
PR 29-JUN-2001; 2001US-302344P.  
XX  
XX (STRD ) UNIV LELAND STANFORD JUNIOR.  
XX  
XX McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;  
PI MPI: 2003-210268/20.  
XX P-PSDB; ABP70434.  
XX  
XX New nucleic acid comprising a mammalian T cell immunoglobulin domain  
PT and Mucin domain gene sequences, useful for treating cancer or asthma,  
XX allergy, eczema or autoimmune disease -  
XX  
XX  
XX Claim 5; Page 76; 94pp; English.  
XX  
XX The present sequence encodes a murine T cell immunoglobulin domain and  
CC mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1,  
CC TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with  
CC conserved IgV and mucin domains. The locus comprising the TIM family is  
CC genetically associated with immune dysfunction, including asthma. The  
CC TIM gene family is located within a region of human chromosome 5 that  
CC is commonly deleted in malignancies and myelodysplastic syndrome.  
CC Variants of TIM-1 and TIM-3 are associated with susceptibility to  
CC airway hyperreactivity and allergic T cell responses, and other variants  
CC associated with protection against these responses. T cells express TIM  
CC proteins, which critically regulate CD4 T cell differentiation. Th1  
CC cells preferentially express TIM-3, while Th2 cells preferentially  
CC express TIM-1. TIM polypeptides and polynucleotides are useful for  
CC treating cancer, asthma, allergies, eczema or autoimmune diseases.  
XX  
XX Sequence 2725 BP; 805 A; 587 C; 641 G; 692 T; 0 other;  
SQ

Query Match 96.8%; Score 2624.6; DB 25; Length 2725;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 2655; Conservative 6; Mismatches 20; Indels 2; Gaps 2;

QY 18 GGAATTCCTCCCTCCAGTACTGATGTTTCAGGCTTACCTCAAGTGTCTGCTGCT 77  
DB 45 GGAATTCCTCCCTCCAGTACTGATGTTTCAGGCTTACCTCAAGTGTCTGCTGCT 104

QY 78 GCTGCAACTACTACTGTCAGAGGTCAATGGAAGTGTATTAAGTTGAAGTTGTAATAA 137  
DB 105 GCTGCAACTACTACTGTCAGAGGTCAATGGAAGTGTATTAAGTTGAAGTTGTAATAA 164

QY 138 TGGCTATTCGCTCCGTCAGATTAACCTTACCTATCTGGGACCTGTGCTATGTGCTG 197  
DB 165 TGGCTATTCGCTCCGTCAGATTAACCTTACCTATCTGGGACCTGTGCTATGTGCTG 224

QY 198 GGGCAAGGAGTGTGCTGCTGTCAGAGTGTACCAATGAGTGTCTGAGAACTGATGAAG 257  
DB 225 GGGCAAGGAGTGTGCTGCTGTCAGAGTGTACCAATGAGTGTCTGAGAACTGATGAAG 284

QY 258 AATGTGACATATCAGAAATCCAGAGATACCAAGTAAAGGGGATCTCAACAAAGAGA 317  
DB 285 AATGTGACATATCAGAAATCCAGAGATACCAAGTAAAGGGGATCTCAACAAAGAGA 344

QY 318 TGTGTCTGATCATTAAGAAATGTGACTCTGATGACCAATGGACCTACTGCTGACAGAT 377  
DB 345 TGTGTCTGATCATTAAGAAATGTGACTCTGATGACCAATGGACCTACTGCTGACAGAT 404

QY 378 ACAGTTCCTGCTCTTAATGAATGAATAAATTAAGACCTGAATTAAGATCAAAAGCAC 437  
DB 405 ACAGTTCCTGCTCTTAATGAATGAATAAATTAAGACCTGAATTAAGATCAAAAGCAC 464

QY 438 CAAGTCACTCCAGTCCAGACTGACCTCCATGGGACTCTACTAGAGCTTCTCAAGAACCT 497  
DB 465 CAAGTCACTCCAGTCCAGACTGACCTCCATGGGACTCTACTAGAGCTTCTCAAGAACCT 524

QY 498 AACCAAGGAGAAATGTTTCAGAGACACAGACACTGTGACCTTCATTAATPAACATGG 557  
DB 525 AACCAAGGAGAAATGTTTCAGAGACACAGACACTGTGACCTTCATTAATPAACATGG 584

QY 558 AACCAAAATTTCCAGATGGGCTGATGAATTAAGGACTGTGAGAAAGATGAGACTGC 617  
DB 585 AACCAAAATTTCCAGATGGGCTGATGAATTAAGGACTGTGAGAAAGATGAGACTGC 644

QY 618 TATCCACATTTGAGAGTGGAGTCTGCTGGTGGACCTGGACCTTATCATTTGCTGCT 677  
DB 645 TATCCACATTTGAGAGTGGAGTCTGCTGGTGGACCTGGACCTTATCATTTGCTGCT 704

QY 678 AATCTTAAATGCTATTCCTGTAAGAAAAGATTAATGAGTTGAGCTTATTAACCT 737  
DB 705 AATCTTAAATGCTATTCCTGTAAGAAAAGATTAATGAGTTGAGCTTATTAACCT 764

QY 738 GGCCAACTGCTCCAGAGGGGTGGCAAAATCAGAGACATCAGATTCGGCTGAGGA 797  
DB 765 GGCCAACTGCTCCAGAGGGGTGGCAAAATCAGAGACATCAGATTCGGCTGAGGA 824

QY 798 AATATCTACACCATCAGAGAGAGAACTATATAGTGAAGTGAATTCAAATGACTACTG 857  
DB 825 AATATCTACACCATCAGAGAGAGAACTATATAGTGAAGTGAATTCAAATGACTACTG 884

QY 858 CTAAGTCAACAGCCAGCCATCTGACCCGCTCTGACCTGCCATTTAAAGCTGCC 917  
DB 885 CTAAGTCAACAGCCAGCCATCTGACCCGCTCTGACCTGCCATTTAAAGCTGCC 944

QY 918 CTTCAATTTGCTGCTTGGATTTCCCTTTGGAAAACATAGTGAATGCTGAGCA 977  
DB 945 CTTCAATTTGCTGCTTGGATTTCCCTTTGGAAAACATAGTGAATGCTGAGCA 1003

QY 978 CCTCATTTGAGGTTCTGACACAGCCACTGAGAAAAGATTCAGTTTCTGGGATTAAT 1037  
DB 1004 CCTCATTTGAGGTTCTGACACAGCCACTGAGAAAAGATTCAGTTTCTGGGATTAAT 1063

QY 1038 TAACCTCAAGGGGATTCGACTGTAACTCATGCTATTTGAATGCTCATTTTATCCCT 1097  
DB 1064 TAACCTCAAGGGGATTCGACTGTAACTCATGCTATTTGAATGCTCATTTTATCCCT 1123

QY 1098 GAGTTTCAGGATTCGATTCCTCCACTCCAGAGACTTCATCATGTGCTGTAAGCTCACT 1157  
DB 1124 GAGTTTCAGGATTCGATTCCTCCACTCCAGAGACTTCATCATGTGCTGTAAGCTCACT 1183

QY 1158 CGTGCTTTCATACATTAAGAAATGTTAGTGATGCTTTTGAACATAGAGGTTTGCT 1217  
DB 1184 CGTGCTTTCATACATTAAGAAATGTTAGTGATGCTTTTGAACATAGAGGTTTGCT 1243

QY 1218 AATATCCGAAAGCTCTCTGAACAGTAAAGGGGAAATAAGGCTTAATAGAAAGTGGCT 1277  
DB 1244 AATATCCGAAAGCTCTCTGAACAGTAAAGGGGAAATAAGGCTTAATAGAAAGTGGCT 1303

QY 1278 CTTTGTGATGTTGAAAATCTTAAGAAAGTGTAGCTTTTCTAGAGATTTCTGACCT 1337  
DB 1304 CTTTGTGATGTTGAAAATCTTAAGAAAGTGTAGCTTTTCTAGAGATTTCTGACCT 1362

QY 1338 GAAAGATTAAGAAAAGCCAGGTGACATATGCTTAACAGATATACTTGGGAACCTTAG 1397  
DB 1363 GAAAGATTAAGAAAAGCCAGGTGACATATGCTTAACAGATATACTTGGGAACCTTAG 1422

QY 1398 GCAGAGGGGTATAGTTCAAGTTCAGCCAGGGCTATGCTGGTAAAGTCTCTCAMCMTTC 1457  
 DB 1423 GCAGAGGGGTATAGTTCAAGTTCAGCCAGGGCTATGCTGGTAAAGTCTCTCAMCMTTC 1482  
 QY 1458 CAAGAAGCAAAATTAACATAGAGACAGAGAGGCTGGAGATGAGCTCGACAGTGAAG 1517  
 DB 1483 CAAGAAGCAAAATTAACATAGAGACAGAGAGGCTGGAGATGAGCTCGACAGTGAAG 1542  
 QY 1518 TGCATTGTATCAAGCAGAGGAATCTATATTATGATCCGATGAGCCCAATGAGAAAGCTA 1577  
 DB 1543 TGCATTGTATCAAGCAGAGGAATCTATATTATGATCCGATGAGCCCAATGAGAAAGCTA 1602  
 QY 1578 GGCCTGTAGAGCATGCTTGTAGACTCAAGAGATGAGAGGTAAGGACAACAGATCCC 1637  
 DB 1603 GGCCTGTAGAGCATGCTTGTAGACTCAAGAGATGAGAGGTAAGGACAACAGATCCC 1662  
 QY 1638 CGGGGCTTGGCTGACAGTCTAGCTTACCTTAAAGTCTGAGTCCAAAGTCCCT 1697  
 DB 1663 CGGGGCTTGGCTGACAGTCTAGCTTACCTTAAAGTCTGAGTCCAAAGTCCCT 1722  
 QY 1698 GCTCAGAGTAAAGTCTGATGATCTGGGCAATGTCATGGGGGTTCTCTCTCTCT 1757  
 DB 1723 GCTCAGAGTAAAGTCTGATGATCTGGGCAATGTCATGGGGGTTCTCTCTCTCT 1782  
 QY 1758 CAGAAGACATGACATGACCTGACACACACACACACACACACACACACACACAC 1817  
 DB 1783 CAGAAGACATGACATGACCTGACACACACACACACACACACACACACACACAC 1842  
 QY 1818 ACACACACACATGAAATGAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1877  
 DB 1843 ACACACACACATGAAATGAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1902  
 QY 1878 CTCTACAGAGCT 1937  
 DB 1903 CTCTACAGAGCT 1962  
 QY 1938 AGTAATTTATTCACACACTCAGAGGCTGAGACAGAGGCTGAGAGTTCAGAGACACTG 1997  
 DB 1963 AGTAATTTATTCACACACTCAGAGGCTGAGACAGAGGCTGAGAGTTCAGAGACACTG 2022  
 QY 1998 TGCCCAACACCTGCGACACTCTTCTTACACAGAAAAAGTTACCGCAAGCCTGCTG 2057  
 DB 2023 TGCCCAACACCTGCGACACTCTTCTTACACAGAAAAAGTTACCGCAAGCCTGCTG 2082  
 QY 2058 TCTGTAAAGAAACCTCTGCGAAAGCAAACTTTGACTGTGTGTCTCAAGGGGAACCTG 2117  
 DB 2083 TCTGTAAAGAAACCTCTGCGAAAGCAAACTTTGACTGTGTGTCTCAAGGGGAACCTG 2142  
 QY 2118 ACTCAGACAACTTCTCATTCTCTGAGGAAACTGAGCTGTTCTGACAGAAACAAC 2177  
 DB 2143 ACTCAGACAACTTCTCATTCTCTGAGGAAACTGAGCTGTTCTGACAGAAACAAC 2202  
 QY 2178 GGTGACTGGGACATACGAAGGAGAGCTTCTGACAGCAATCTATAGTACGAAATAT 2237  
 DB 2203 GGTGACTGGGACATACGAAGGAGAGCTTCTGACAGCAATCTATAGTACGAAATAT 2262  
 QY 2238 CTTTGGGAGGACATGCTGACCAAAATGATTTCCAAAGCCGCTGACCTTCACTGCTG 2297  
 DB 2263 CTTTGGGAGGACATGCTGACCAAAATGATTTCCAAAGCCGCTGACCTTCACTGCTG 2322  
 QY 2298 GCTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2357  
 DB 2323 GCTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2382  
 QY 2358 AATTAATTAAGACCCCGATGAAATTAATTAAGTGAAGCAAGAGTACTTGTCTCAAGAT 2417  
 DB 2383 AATTAATTAAGACCCCGATGAAATTAATTAAGTGAAGCAAGAGTACTTGTCTCAAGAT 2442  
 QY 2418 TTTTGTGCTTGGGAGCACTGTGTACATCAGAGACATCTGTATAGTGAAGCAACCA 2477  
 DB 2443 TTTTGTGCTTGGGAGCACTGTGTACATCAGAGACATCTGTATAGTGAAGCAACCA 2502

QY 2478 ACTGTGTACCGTTTTCATGATGATGATTTGTTAGTTGCTTACCTAGCTG 2537  
 DB 2503 ACTGTGTACCGTTTTCATGATGATGATTTGTTAGTTGCTTACCTAGCTG 2562  
 QY 2538 TGGAGTCTGCTGCTTCTTCTAGTGGTATGAGAGGACCATCTAACAAATCCATTAG 2597  
 DB 2563 TGGAGTCTGCTGCTTCTTCTAGTGGTATGAGAGGACCATCTAACAAATCCATTAG 2622  
 QY 2598 AGATTAACAGCTCTCATGACAGAGGAAACCTATCTCAATGTTTAAAGTAAACT 2657  
 DB 2623 AGATTAACAGCTCTCATGACAGAGGAAACCTATCTCAATGTTTAAAGTAAACT 2682  
 QY 2658 GTACTGGCAAGTACTTTCAGCATTAACCAATTAACCAATTAACCAATTAACCA 2700  
 DB 2683 GTACTGGCAAGTACTTTCAGCATTAACCAATTAACCAATTAACCAATTAACCA 2725  
 RESULT 10  
 AB268329  
 ID AB268329 standard; DNA; 862 BP.  
 XX  
 AC AB268329;  
 XX  
 DT 22-APR-2003 (first entry)  
 XX  
 XX  
 DE Nucleotide sequence of murine TIM-3 ES-HBA and DBA/2J allele.  
 XX  
 KM T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;  
 KM TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;  
 KM myeloid/plastic syndrome; airway hyperreactivity; cancer; asthma;  
 KM allergic T cell response; autoimmune disease; gene; ss.  
 OS Mus musculus.  
 FH Key Location/Qualifiers  
 FT CDS 17..862  
 FT /\*tag= a  
 FT /product= "TIM-3"  
 PN MO2003002722-A2.  
 PD 09-JAN-2003.  
 PF 01-JUL-2002; 2002MO-US20890.  
 PR 29-JUN-2001; 2001US-302344P.  
 PA (STRD ) UNIV LEIAND STANFORD JUNIOR.  
 XX  
 XX McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GD, Kuchroo V;  
 DR MPI. 2003-210268/20.  
 DR P-PSDB; ABP70435.  
 PT New nucleic acid comprising a mammalian T cell immunoglobulin domain  
 PT and mucin domain gene sequences, useful for treating cancer or asthma,  
 PT allergy, eczema or autoimmune disease -  
 XX  
 PS Claim 5, Page 77, 94pp; English.  
 CC The present sequence encodes a murine T cell immunoglobulin domain and  
 CC mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1,  
 CC TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with  
 CC conserved Igv and mucin domains. The locus comprising the TIM family is  
 CC genetically associated with immune dysfunction, including asthma. The  
 CC TIM gene family is located within a region of human chromosome 5 that  
 CC is commonly deleted in malignancies and myelodysplastic syndrome.  
 CC Variants of TIM-1 and TIM-3 are associated with susceptibility to  
 CC airway hyperreactivity and allergic T cell responses. T cells express  
 CC associated with protection against these responses. T cells express TIM  
 CC proteins, which critically regulate CD4 T cell differentiation. Th1  
 CC cells preferentially express TIM-3, while Th2 cells preferentially  
 CC express TIM-1. TIM polypeptides and polynucleotides are useful for

CC treating cancer, asthma, allergies, eczema or autoimmune diseases.

XX Sequence 862 BP; 254 A; 197 C; 198 G; 213 T; 0 other;

SQ

Query Match 31.2%; Score 844.4; DB 25; Length 862;

Best Local Similarity 98.7%; Pred. No. 1.6e-227;

Matches 851; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

24 CCCCTCCCAAGTACATGTTTCAAGTCTTACCTCACTGCTGCTGCTGCTGCA 83

1 CCCCTCCCAAGTACATGTTTCAAGTCTTACCTCACTGCTGCTGCTGCTGCA 60

84 ACTACTACTGCAAGGTCATTTGAAGTGTAAAGTGTGTTAAAGTGTAAAGTGT 143

61 ACTACTACTGCAAGGTCATTTGAAGTGTAAAGTGTGTTAAAGTGTAAAGTGT 120

144 TCTGCCCTGAGTATACCTCTACCTACATCTGGAGACCTGTGCTGCTGCTGCA 203

121 TCTGCCCTGAGTATACCTCTACCTACATCTGGAGACCTGTGCTGCTGCTGCA 180

204 GGGATTTCTGCTTGTGTCAGAGTGTACCAATGAGTGTCTGAGAACTGATGAAAGT 263

181 GGGATTTCTGCTTGTGTCAGAGTGTACCAATGAGTGTCTGAGAACTGATGAAAGT 240

264 GACATATCAGAAATCCAGAGATACCGCTAAAGGGGATCTCAACAAAGAGATGTC 323

241 GACATATCAGAAATCCAGAGATACCGCTAAAGGGGATCTCAACAAAGAGATGTC 300

324 TCTGATCATTAAGAAATGTGACTCTGATGACCAATGGAGACCTGCTGCTGCTGCA 383

301 TCTGATCATTAAGAAATGTGACTCTGATGACCAATGGAGACCTGCTGCTGCTGCA 360

384 CCTGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443

361 CCTGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420

444 CACTCCAGCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 503

421 CACTCCAGCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

504 GAGAGAAATGTTTCAAGAGACACAGACATGCTGCTGCTGCTGCTGCTGCTGCTG 563

481 GAGAGAAATGTTTCAAGAGACACAGACATGCTGCTGCTGCTGCTGCTGCTGCTG 540

564 AATTTCACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 623

541 AATTTCACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 600

624 CATTGAGTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 683

601 CATTGAGTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

684 TAAATGCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743

661 TAAATGCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720

744 CTGCTCTTCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 803

721 CTGCTCTTCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780

804 CTACACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 863

781 CTACACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840

864 CAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 885

841 CAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 862

RESULT 11

AAA51918

AAA51918 standard; DNA; 843 BP.

AC AAA51918;

DT 31-OCT-2000 (first entry)

XX T helper cell differentially expressed gene.

XX T helper cell; differential expression; 102 gene; immunomodulator;

XX anti-inflammatory; anti-arthritic; antibacterial; immunosuppressive;

XX chymotrypsin; anti-thyroid; anti-asthmatic; anti-allergic; antiviral;

XX protozoacide; lymphocyte; modulator; gene therapy; ss.

XX Mus sp.

XX US6084083-A.

XX 04-JUL-2000.

XX 28-MAR-1997; 97US-0829525.

XX 01-MAR-1996; 96US-0609583.

XX 03-MAR-1995; 95US-0398633.

XX 07-JUN-1995; 95US-0487748.

XX (MILL-) MILLENNIUM PHARM INC.

XX Levinson DA;

XX WPI; 2000-464385/40.

XX New isolated human 200 gene products or polypeptides, useful for

XX treating and diagnosing immune disorders, especially T helper

XX lymphocyte-related disorders

XX Disclosure; Column 121-122; 107pp; English.

XX Genes which are differentially expressed within and among T helper (TH)

XX cells and TH cell subpopulations, e.g. TH0, TH1 and TH2 subpopulations,

XX can be used diagnostically or as targets for therapeutic intervention.

XX The polypeptides are useful for treating and diagnosing of immune

XX disorders, especially T lymphocyte-related disorders. These disorders

XX include chronic inflammatory diseases and disorders (e.g. Crohn's

XX disease, reactive arthritis, Lyme disease, Hashimoto's thyroiditis or

XX Grave's disease), or atopic conditions (e.g. asthma and allergy,

XX including allergic rhinitis or food allergies). Also included are

XX certain pathogen susceptibilities (e.g. leishmaniasis), and viral (e.g.

XX HIV) or bacterial (e.g. tuberculosis or leprosy) infections.

XX Sequence 843 BP; 253 A; 185 C; 196 G; 209 T; 0 other;

Query Match 31.1%; Score 843; DB 21; Length 843;

Best Local Similarity 100.0%; Pred. No. 3.8e-227;

Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

40 ATGTTTCAAGTCTTACCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 99

1 ATGTTTCAAGTCTTACCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60

100 TCATTGGAAGATGTTTAAAGTGTGAGGTTGTTAAATGCTTATGCTGCTGCACTTAC 159

61 TCATTGGAAGATGTTTAAAGTGTGAGGTTGTTAAATGCTTATGCTGCTGCACTTAC 120

160 ACTCTACCTTACATCTGAGACATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 219

121 ACTCTACCTTACATCTGAGACATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

220 TCACAGTGTACCAATGATGTTGCTCAGAACTGATGAAAGAAATGTGACATATCAGAAATCC 279

181 TCACAGTGTACCAATGATGTTGCTCAGAACTGATGAAAGAAATGTGACATATCAGAAATCC 240

280 AGCAGATACCAAGCTAAAGGCGATCTCAACAAAGAGATGTGCTCTGATCATTAAGAT 339

241 AGCAGATACCAAGCTAAAGGCGATCTCAACAAAGAGATGTGCTCTGATCATTAAGAT 300

```

QY 340 GTGACTGTGATGACCATGGAGCACTTACTGTGTCAGAGATACAGTTCCTGTCTTATGAT 399
XX |||
DB 301 GTGACTGTGATGACCATGGAGCACTTACTGTGTCAGAGATACAGTTCCTGTCTTATGAT 360
QY 400 GATTAATAATTTGAATGAAATTTAGACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 459
XX |||
DB 361 GATTAATAATTTGAATGAAATTTAGACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 460 GCCCATGGGGAGCTTACTTACAGCTTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 519
XX |||
DB 421 GCCCATGGGGAGCTTACTTACAGCTTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 520 GAGACAGACAGACAGCTGTGACCTTCCATATATACAAATGAGAGAGAGAGAGAGAGAGAG 579
XX |||
DB 481 GAGACAGACAGACAGCTGTGACCTTCCATATATACAAATGAGAGAGAGAGAGAGAGAGAG 540
QY 580 GATGAATTTAGAGACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 639
XX |||
DB 541 GATGAATTTAGAGACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 640 TCTGCTGGGGTGTACCTGTGACCTTATTCATGTTGTTCTTAACTTAAATGATTCCTGT 699
XX |||
DB 601 TCTGCTGGGGTGTACCTGTGACCTTATTCATGTTGTTCTTAACTTAAATGATTCCTGT 660
QY 700 AAGAAAAAGAACTTATTCAGATTGAGCTTATTAACACTGAGAGAGAGAGAGAGAGAGAG 759
XX |||
DB 661 AAGAAAAAGAACTTATTCAGATTGAGCTTATTAACACTGAGAGAGAGAGAGAGAGAGAG 720
QY 760 TTGGCAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 819
XX |||
DB 721 TTGGCAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 820 AACGTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 879
XX |||
DB 781 AACGTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 880 TCC 882
XX |||
DB 841 TCC 843

```

RESULT 12  
AA170255  
ID AA170255 standard; cDNA; 843 BP.

```

XX AC AA170255;
XX DT 07-JAN-2002 (first entry)
XX DE Mouse 200 gene, preferentially expressed in TH1 cells.
XX KM Mouse; 200 gene; T helper; T lymphocyte; T cell; TH1;
XX KM receptor; differential expression; immune disorder; psoriasis;
XX KM multiple sclerosis; insulin-dependent diabetes; antidiabetic;
XX KM antipsoriatic; diagnosis; therapy; ss.
XX OS Mus musculus.
XX PN US6288218-B1.
XX PD 11-SEP-2001.
XX PF 25-SEP-1997; 97US-0937399.
XX PR 01-MAR-1996; 96US-0609583.
XX PR 03-MAR-1995; 95US-0398633.
XX PR 07-JUN-1995; 95US-0487748.
XX PA (LEVI/) LEVINSON D A.
XX PI Levinson DA;
XX WP1; 2001-647189/74.

```

```

DR P-PSDB; AAM50225.
XX
XX Detecting 200 gene expression in a sample, useful for treating and
PT diagnosing immune disorders, especially T lymphocyte-related disorders.
PT comprises detecting the presence of a 200 gene product or an RNA
PT encoding the 200 gene product.
XX
PS Disclosure; Column 123-124; 108pp; English.
XX
XX The present sequence is that of cDNA comprising the complete open
CC reading frame of the mouse 200 gene. It encodes a 281-amino acid
CC protein (see AAM50225). The cDNA was isolated using the 'T cell
CC clone paradigm' and the 'transgenic T cell paradigm' approaches
CC to the identification of genes that are differentially expressed
CC in helper T cell (TH) subpopulations. The gene is expressed at
CC a higher level in fully differentiated TH1 than in TH2 cells. It
CC can be used diagnostically or as a target for therapeutic
CC intervention for the treatment of immune disorders. A claimed
CC method for diagnosing a TH cell subpopulation-related immune
CC disorder involves detecting the level of a human 200 gene (see
CC AAI70254) product, or an RNA encoding it, so that if the level
CC differs from that in a control sample, the disorder is diagnosed.
CC The disorder is especially a TH1 cell subpopulation-related immune
CC disorder, such as multiple sclerosis, psoriasis or insulin-dependent
CC diabetes (claimed). A method for detecting murine 200 gene
CC expression is also claimed. In addition to the 200 gene, the
CC invention provides other genes that are differentially expressed
CC within and among TH cells and TH cell subpopulations and which can be
CC used in methods for the diagnosis, prognosis, evaluation and treatment
CC of TH cell subpopulation-related disorders, for the identification of
CC subjects exhibiting a predisposition to such conditions, for
CC monitoring patients undergoing clinical evaluation for the
CC treatment of such disorders, and for monitoring the efficacy of
CC compounds used in clinical trials. Other immune disorders that can
CC be treated/diagnosed include Crohn's disease, reactive arthritis,
CC Lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis,
CC Grave's disease, contact dermatitis, graft rejection, giant verus
CC host disease, sarcoidosis, atopic conditions, asthma, allergy,
CC allergic rhinitis, food allergy, eosinophilia, conjunctivitis,
CC glomerular nephritis, helminthic infection (e.g. leishmaniasis),
CC viral infection (e.g. HIV), and bacterial infection (e.g.
CC tuberculosis and leprosy).
XX
SQ Sequence 843 BP; 253 A; 185 C; 196 G; 209 T; 0 other;

```

Query Match 31.1%; Score 843; DB 22; Length 843;  
Best Local Similarity 100.0%; Pred. No. 3; Be-227;  
Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps

```

QY 40 ATGTTTCAAGTCTTACCTCAACTGTGCTGCTGCTGCACTACTACTTGCAGAG 99
XX |||
DB 1 ATGTTTCAAGTCTTACCTCAACTGTGCTGCTGCTGCTGCACTACTACTTGCAGAG 60
QY 100 TCATTGAAGATGTTTAAAGTTAGAGTTGAGTTGTAATAATGCTATTCGCCCTGACGTTAC 159
XX |||
DB 61 TCATTGAAGATGTTTAAAGTTAGAGTTGAGTTGTAATAATGCTATTCGCCCTGACGTTAC 120
QY 160 ACTTACCTCACTCTGAGACACTTGTGCTATGTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 219
XX |||
DB 121 ACTTACCTCACTCTGAGACACTTGTGCTATGTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 220 TCACAGTGAACAATGATGATGCTGAGAACTGATGAAGAAATGATGATCAGAAATCC 279
XX |||
DB 181 TCACAGTGAACAATGATGATGCTGAGAACTGATGAAGAAATGATGATCAGAAATCC 240
QY 280 AGCAGATACCAAGCTTAAAGGGGAGATCTCAACAAAGAGATGTGTCTGTGATCAATAAAGAT 339
XX |||
DB 241 AGCAGATACCAAGCTTAAAGGGGAGATCTCAACAAAGAGATGTGTCTGTGATCAATAAAGAT 300
QY 340 GTGACTGTGATGACCATGGAGCACTTACTGTGTCAGAGATACAGTTCCTGTCTTATGAT 399
XX |||
DB 301 GTGACTGTGATGACCATGGAGCACTTACTGTGTCAGAGATACAGTTCCTGTCTTATGAT 360

```



[illegible]

RESULT 13
AAF82615
ID AAF82615 standard; cDNA; 843 BP.
XX
AC AAF82615;
XX
DT 18-JUN-2001 (first entry)
DE Murine TH1 specific 200 gene coding sequence.
XX
KW Murine; T helper cell; TH cell; TH1; TH2; immunomodulator;
KW anti-inflammatory; antiallergic; dermatological; antiviral
KW antibacterial; T helper lymphocyte modulator; gene therapy
KW TH specific gene; 200 gene; immune disorder; inflammation;
infection; ss.
XX
OS Mus sp.
XX
FH Location/Qualifiers
FT 1..843
CDS /tag= a
FT /product= "200 gene product"
FT
XX
PN US6204371-B1.
PD
20-MAR-2001.
PF 01-MAR-1996; 96US-0609583.
XX
PR 03-MAR-1995; 95US-0398633.
PR 07-JUN-1995; 95US-0487748.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Levinson DA;
XX

XX	WPI, 2001-272703/28.
DX	
PT	New murine or human 200 genes and their corresponding polypeptides,
PT	useful for treating or diagnosing immune disorders, especially T helper
PT	lymphocyte-related disorders, e.g. inflammatory diseases or allergies
PT	-
PS	
XX	Claim 6; Fig 17, 109pp; English.
XX	
CC	The present sequence is expressed at higher levels in T helper 1 (TH1)
CC	cells than in T helper 2 (TH2) cells. The invention relates to an
CC	isolated nucleic acid molecule, which comprises the full length murine
CC	200 gene or full length human 200 gene nucleotide sequence. The nucleic
CC	acids are useful for treating or diagnosing immune disorders.
CC	especially T helper lymphocyte-related disorders, e.g. inflammatory
CC	diseases (e.g. Crohn's disease), multiple sclerosis, Grave's disease,
CC	contact dermatitis, psoriasis, asthma and allergies, or certain viral
CC	(e.g. HIV) or bacterial (e.g. tuberculosis) infections.
XX	
SQ	Sequence 843 BP; 253 A; 185 C; 196 G; 209 T; 0 other;

Query Match	Similarity	3.1.%	Score 843	DB 22	Length 843
Best Local	Similarity	100.0%	Pred. No. 3.8e-22		
Matches 843	Conservative	0	Mismatches 0	Indels 0	Gaps 0
QY	40	ATGTTTCAGGCTCTTACCCCTCACTGTGTCTGCTGCTGCTCAACTACTACTTGCAGG	99		
Db	1	ATGTTTTCAGGCTCTTACCCCTCACTGTGTCTGCTGCTGCTCAACTACTACTTGCAGG	60		
QY	100	TCATTGGAAGATGTTATTAAGGTTAGGTTGGTAAATAATGCTTATCTGCCCTGCAGTTC	159		
Db	61	TCATTGGAAGATGTTATTAAGGTTAGGTTGGTAAATAATGCTTATCTGCCCTGCAGTTC	120		
QY	160	ACTTACCTACATCTGGGACACTTGTGCTATGTCTGCGGCAAGGATTTCTGCTTGG	219		
Db	121	ACTTACCTACATCTGGGACACTTGTGCTATGTCTGCGGCAAGGATTTCTGCTTGG	180		
QY	220	TCACAGTGTACCAATGAGTTGTCTCAAACTGATGGAAGAAATGTGACATATCAGAAATCC	279		
Db	181	TCACAGTGTACCAATGAGTTGTCTCAAACTGATGGAAGAAATGTGACATATCAGAAATCC	240		
QY	280	AGCAGATACCCAGCTTAAAGGGGATTTCAACAAGAGATGTGTCTGTATCTAATAAGAT	339		
Db	241	AGCAGATACCCAGCTTAAAGGGGATTTCAACAAGAGATGTGTCTGTATCTAATAAGAT	300		
QY	340	GTGACTCTGGAGTGAACATGGGGACATGCTGTCGAGAGATPACATTCCTGCTTATGAAT	399		
Db	301	GTGACTCTGGAGTGAACATGGGGACATGCTGTCGAGAGATPACATTCCTGCTTATGAAT	360		
QY	400	GATAAAAAATTAGAACTGAAATTTAGACATCAAAAGCAGCCAAAGTCACTTCAGTCCAGACT	459		
Db	361	GATAAAAAATTAGAACTGAAATTTAGACATCAAAAGCAGCCAAAGTCACTTCAGTCCAGACT	420		
QY	460	GCCCATGGGGACTTACTTACAGCTTCTCCAGAACCTTAAACAAGGAGAAATGTTCA	519		
Db	421	GCCCATGGGGACTTACTTACAGCTTCTCCAGAACCTTAAACAAGGAGAAATGTTCA	480		
QY	520	GAGACACAGACACTGGGTGACCCCTGCATATATACATGGAAGAAAAATTTCCATGGGGCT	579		
Db	481	GAGACACAGACACTGGGTGACCCCTGCATATATACATGGAAGAAAAATTTCCATGGGGCT	540		
QY	580	GATCAAAATTAAGACTCTGAGAAACGATCAGAACCTGTATCCACATTGAGTGGAGTTC	639		
Db	541	GATCAAAATTAAGACTCTGAGAAACGATCAGAACCTGTATCCACATTGAGTGGAGTTC	600		
QY	640	TCGTGTGGGTGACCCCTGGCACTTATCATTTGTGTTTATCTTTAATATGTAATTCCTGT	699		
Db	601	TCGTGTGGGTGACCCCTGGCACTTATCATTTGTGTTTATCTTTAATATGTAATTCCTGT	660		
QY	700	AAGAAAAAAGATTTATGAGTTTGAAGCTTATTAATCACTGCGCAACTTGGCTCCAGAGGG	759		
Db	661	AAGAAAAAAGATTTATGAGTTTGAAGCTTATTAATCACTGCGCAACTTGGCTCCAGAGGG	720		







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using BW model

Run on: November 21, 2003, 16:48:30 ; Search time 166.515 Seconds  
(without alignments)  
7183.411 Million cell updates/sec

Title: US-10-004-633-8

Perfect score: 2710

Sequence: 1 ngctgaccacgcgtccgca.....aaaaaaaaagcgccgcgc 2710

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2706.6	99.9	2710	1	US-08-487-748A-8
2	2706.6	99.9	2710	3	US-08-480-070C-8
3	2706.6	99.9	2710	3	US-08-829-525-8
4	2706.6	99.9	2710	3	US-08-609-583A-8
5	2706.6	99.9	2710	3	US-08-937-399-8
6	2706.6	99.9	2710	4	US-09-310-367-8
7	2706.6	99.9	2710	4	US-09-032-337-8
8	2706.6	99.9	2710	4	US-09-464-231-8
9	843	31.1	843	3	US-08-829-525-22
10	843	31.1	843	3	US-08-609-583A-23
11	843	31.1	843	3	US-08-937-399-22
12	843	31.1	843	4	US-09-310-367-22
13	843	31.1	843	4	US-09-032-337-22
14	843	31.1	843	4	US-09-464-231-22
15	499.8	18.4	2236	3	US-08-829-525-23
16	499.8	18.4	2236	3	US-08-609-583A-23
17	499.8	18.4	2236	3	US-08-937-399-23
18	499.8	18.4	2236	4	US-09-310-367-23
19	499.8	18.4	2236	4	US-09-032-337-23
20	499.8	18.4	2236	4	US-09-464-231-23
21	456	16.8	903	3	US-08-829-525-37
22	456	16.8	903	3	US-08-609-583A-37
23	456	16.8	903	3	US-08-937-399-37
24	456	16.8	903	4	US-09-310-367-37
25	456	16.8	903	4	US-09-032-337-37
26	456	16.8	903	4	US-09-464-231-37
27	240.8	8.9	313	3	US-08-398-633-8

28	64.4	2.4	1380	4	US-09-620-312D-59	Sequence 59, Appl
29	62.8	2.3	7218	1	US-08-233-463-14	Sequence 14, Appl
30	62	2.3	2093	1	US-08-287-001A-1	Sequence 1, Appl
31	62	2.3	2093	5	PCT-US95-09941-1	Sequence 1, Appl
32	60.8	2.2	2310	3	US-08-458-434A-1	Sequence 1, Appl
33	60.8	2.2	9299	3	US-08-458-434A-7	Sequence 7, Appl
34	60.4	2.2	80595	3	US-09-078-294-3	Sequence 3, Appl
35	59	2.2	538	4	US-09-495-050A-105	Sequence 105, App
36	58.8	2.2	80246	3	US-09-078-294-4	Sequence 4, Appl
37	57.6	2.1	3358	3	US-08-957-063-4	Sequence 4, Appl
38	57.6	2.1	3358	4	US-09-487-685-4	Sequence 4, Appl
39	57.6	2.1	3358	4	US-08-802-805D-4	Sequence 4, Appl
40	56	2.1	3381	3	US-09-009-119-1	Sequence 1, Appl
41	56	2.1	3381	4	US-09-371-507-1	Sequence 1, Appl
42	56	2.1	3383	5	PCT-US95-09098-1	Sequence 1, Appl
43	55.4	2.0	46718	4	US-09-81C-093-3	Sequence 3, Appl
44	55.2	2.0	200	4	US-09-354-147C-31	Sequence 31, Appl
45	55.2	2.0	564	1	US-08-117-362-32	Sequence 32, Appl

#### ALIGNMENTS

RESULT 1  
US-08-487-748A-8  
; Sequence 8, Application US/08487748A  
; Patent No. 5721351  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITILE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487, 748A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO. 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2710 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 40..885  
; US-08-487-748A-8  
Query Match 99.9%; Score 2706.6; DB 1; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 GTCCAGCCACCGTCGCGATTTCCTCCCAAGTACTCATGTTCACAGCTTACCCATCA 61

Db 2 GTGACCCACGCGTCCGGATTTCCCTCCCAAGTACTGATGTTTCAAGGCTTAACTTCA 61  
Qy 62 ACTGTGCTGTGCTGTGCTGCACTAATTGCAAGGTCAATTGGAAGTGTATTAAG 121  
Db 62 ACTGTGCTGTGCTGTGCTGCACTAATTGCAAGGTCAATTGGAAGTGTATTAAG 121  
Qy 122 TTGAGGTGTGTAATAATGCTATCTGCTGCGATTAACCTTAACCTTAACCTGGAAC 181  
Db 122 TTGAGGTGTGTAATAATGCTATCTGCTGCGATTAACCTTAACCTTAACCTGGAAC 181  
Qy 182 TTGTGCTATGTGCTGGGCAAGGATTTCTGCTTGTGCTACAGTGAACAATGAGTTG 241  
Db 182 TTGTGCTATGTGCTGGGCAAGGATTTCTGCTTGTGCTACAGTGAACAATGAGTTG 241  
Qy 242 TCGAAGCTGATGAATAATGTGATCATTCAGAAATCCAGCAGATACAGCTTAAGGGG 301  
Db 242 TCGAAGCTGATGAATAATGTGATCATTCAGAAATCCAGCAGATACAGCTTAAGGGG 301  
Qy 302 ATCTCAACAAGAGATGTGTCTGTATCATTAAGATGTGATCTTGATGACATGGGA 361  
Db 302 ATCTCAACAAGAGATGTGTCTGTATCATTAAGATGTGATCTTGATGACATGGGA 361  
Qy 362 CCTACTGCTGACAGATACAGTTCCCTGCTTATGATGATTAATAATTAAGACTGAAT 421  
Db 362 CCTACTGCTGACAGATACAGTTCCCTGCTTATGATGATTAATAATTAAGACTGAAT 421  
Qy 422 TGAACATCAAGAGCAGAGTCACTCCAGCTCAGCTGAGCTGAGGCTTACTTACAG 481  
Db 422 TGAACATCAAGAGCAGAGTCACTCCAGCTCAGCTGAGCTGAGGCTTACTTACAG 481  
Qy 482 CTCTTCCAAAGAACCTTACACGAGAGAAATGTTCAAGACACAGACACTGTGTACC 541  
Db 482 CTCTTCCAAAGAACCTTACACGAGAGAAATGTTCAAGACACAGACACTGTGTACC 541  
Qy 482 CTCTTCCAAAGAACCTTACACGAGAGAAATGTTCAAGACACAGACACTGTGTACC 541  
Db 482 CTCTTCCAAAGAACCTTACACGAGAGAAATGTTCAAGACACAGACACTGTGTACC 541  
Qy 542 TCCATATATACATGAGCAAAAATTTCCATGAGGCTGATGAATTAAGACTTGAAG 601  
Db 542 TCCATATATACATGAGCAAAAATTTCCATGAGGCTGATGAATTAAGACTTGAAG 601  
Qy 602 AAACGATCAGAACTGCTATCCATTTGAGAGTGGAGTCTGCTGGGTTGACCTGGCAC 661  
Db 602 AAACGATCAGAACTGCTATCCATTTGAGAGTGGAGTCTGCTGGGTTGACCTGGCAC 661  
Qy 662 TTAATCATGTGTCTTAATTCCTTAATGTAATTCCTGTAAGAAAGAAAGTTATCGAGT 721  
Db 662 TTAATCATGTGTCTTAATTCCTTAATGTAATTCCTGTAAGAAAGAAAGTTATCGAGT 721  
Qy 722 TGAAGCTTATTAACACTGAGCACTTGTCTCCAGAGGGTTGGCAAAATGACAGAGTCA 781  
Db 722 TGAAGCTTATTAACACTGAGCACTTGTCTCCAGAGGGTTGGCAAAATGACAGAGTCA 781  
Qy 782 GGAATTCGCTCTGAGAAATATATCTACACATCGAGAGAAACGTATATGAAGTGAAGT 841  
Db 782 GGAATTCGCTCTGAGAAATATATCTACACATCGAGAGAAACGTATATGAAGTGAAGT 841  
Qy 842 CAAATAGATATCTGTGATAGTCAACAGAGCAAGCATCTGACCCCTGTGAGTGGC 901  
Db 842 CAAATAGATATCTGTGATAGTCAACAGAGCAAGCATCTGACCCCTGTGAGTGGC 901  
Qy 902 ACTTTAAAGGCTCGGCTTCAATTTCTGACTTTGATTTCCCTTTTGTGAAACCTATG 961  
Db 902 ACTTTAAAGGCTCGGCTTCAATTTCTGACTTTGATTTCCCTTTTGTGAAACCTATG 961  
Qy 962 ATATGCTACTTGGCAACCTCATTTGAGAGTTCTGACACAGCCACTGAGAAAAGTTTCA 1021  
Db 962 ATATGCTACTTGGCAACCTCATTTGAGAGTTCTGACACAGCCACTGAGAAAAGTTTCA 1021  
Qy 1022 GTTTTCTGGGATTAATTAATCTCAAGAGGATTCGACTGTAATCTGCTACATTTAAT 1081  
Db 1022 GTTTTCTGGGATTAATTAATCTCAAGAGGATTCGACTGTAATCTGCTACATTTAAT 1081  
Qy 1082 GCTCATTTTATCCCTGAGTTTCAAGGATGAGATTCCTCCACCTCAGAGACTTCAATCATG 1141  
Db 1082 GCTCATTTTATCCCTGAGTTTCAAGGATGAGATTCCTCCACCTCAGAGACTTCAATCATG 1141

Db 1082 GCTCATTTTATCCCTGAGTTTCAAGGATGAGATTCCTCCACCTCAGAGACTTCAATCATG 1141  
Qy 1142 CGTGTGAAGCTCACTCGTCTTTTATACATTAAGAAATGTTAGTGTGATCTTTGAGA 1201  
Db 1142 CGTGTGAAGCTCACTCGTCTTTTATACATTAAGAAATGTTAGTGTGATCTTTGAGA 1201  
Qy 1202 CATTAAGGTTGTGTAATTCGCAAGCTCCGTAACAGTGTGGGAAATTAAGGCTTA 1261  
Db 1202 CATTAAGGTTGTGTAATTCGCAAGCTCCGTAACAGTGTGGGAAATTAAGGCTTA 1261  
Qy 1262 GATTAAGAGTGTGCTTTTGTGATGTTGAAATCTTAAGAAATGTTAGTGTGTTCT 1321  
Db 1262 GATTAAGAGTGTGCTTTTGTGATGTTGAAATCTTAAGAAATGTTAGTGTGTTCT 1321  
Qy 1322 AGAATTTCTGACCTTGAAGATTAAGAAAAGCAAGTGGCATATGCTTAACAGATAT 1381  
Db 1322 AGAATTTCTGACCTTGAAGATTAAGAAAAGCAAGTGGCATATGCTTAACAGATAT 1381  
Qy 1382 AACCTGGGAACCTTGAAGGATGATTAAGTTCAAGTCCAGAGGCTTATGCTGTA 1441  
Db 1382 AACCTGGGAACCTTGAAGGATGATTAAGTTCAAGTCCAGAGGCTTATGCTGTA 1441  
Qy 1442 AGACTGTCTCAACATCCAAAAGCAAAATTAACATAGAGACAGAGAGCTGAGATGA 1501  
Db 1442 AGACTGTCTCAACATCCAAAAGCAAAATTAACATAGAGACAGAGAGCTGAGATGA 1501  
Qy 1502 GGCCTGGACAGTGAAGTCAATTTGTATCAAGCAGAGAAATCTATATTTGATCTGAGAC 1561  
Db 1502 GGCCTGGACAGTGAAGTCAATTTGTATCAAGCAGAGAAATCTATATTTGATCTGAGAC 1561  
Qy 1562 CCAATGAAAAGCTAGGCTGTGATGATCTTGTGATCTTCAAGAGATGAGAGGTTA 1621  
Db 1562 CCAATGAAAAGCTAGGCTGTGATGATCTTGTGATCTTCAAGAGATGAGAGGTTA 1621  
Qy 1622 AGGCAACAAGATCCCGGGGCTTGGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1681  
Db 1622 AGGCAACAAGATCCCGGGGCTTGGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1681  
Qy 1682 GTTCAACAAGATCCCTGTCTCAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1741  
Db 1682 GTTCAACAAGATCCCTGTCTCAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1741  
Qy 1742 GGTGTGCTCTCTGCAAGAAAGACATGCAATGCTGCAACACACACACACACAC 1801  
Db 1742 GGTGTGCTCTCTGCAAGAAAGACATGCAATGCTGCAACACACACACACACACAC 1801  
Qy 1802 AC 1861  
Db 1802 AC 1861  
Qy 1862 TCTCTAATAATGATCTCTACAGAGCTCTCTCTGCTCTGTTAAGACATGAGTGGAG 1921  
Db 1862 TCTCTAATAATGATCTCTACAGAGCTCTCTCTGCTCTGTGTAAAGACATGAGTGGAG 1921  
Qy 1922 CATGGAGAGAGTCAAGTAATTTATTCAGAGCTCAAGAAAGCTGGAGAGAGTGA 1981  
Db 1922 CATGGAGAGAGTCAAGTAATTTATTCAGAGCTCAAGAAAGCTGGAGAGAGTGA 1981  
Qy 1982 GAGTTCAGAGAGCTGTGCTCAACACTGCTCAACTCTTTACACAAAGAAAGTTACC 2041  
Db 1982 GAGTTCAGAGAGCTGTGCTCAACACTGCTCAACTCTTTACACAAAGAAAGTTACC 2041  
Qy 2042 CGCAAGAGCTGTGCTGTGTAAAGAAACCTCGAAAGCAAACTTTGACTGTGTG 2101  
Db 2042 CGCAAGAGCTGTGCTGTGTAAAGAAACCTCGAAAGCAAACTTTGACTGTGTG 2101  
Qy 2102 TGCTCAAGGGAATCTGACTCAACAATTTCTCAATTCCTGAGGAAACCTGAGCTGTTTC 2161  
Db 2102 TGCTCAAGGGAATCTGACTCAACAATTTCTCAATTCCTGAGGAAACCTGAGCTGTTTC 2161  
Qy 2162 TGAAGAAAGAAACACCGGTGACTGGGACATAGAAAGGAGAGACTTGTGACGAACTTATA 2221  
Db 2162 TGAAGAAAGAAACACCGGTGACTGGGACATAGAAAGGAGAGACTTGTGACGAACTTATA 2221

QY 2222 TAGTCAGCAAAATATCTTTGGAGAGACGTGTCACCAATGTGATTTCCAAGCCGGTGG 2281  
DB 2222 TAGTCAGCAAAATATCTTTGGAGAGACGTGTCACCAATGTGATTTCCAAGCCGGTGG 2281  
QY 2282 ACCTCAGTTTCAATCTGGCTTACAGCTGCTGCGCCAGTGCCTTGATCTGTGTGGCTCC 2341  
DB 2282 ACCTCAGTTTCAATCTGGCTTACAGCTGCTGCGCCAGTGCCTTGATCTGTGTGGCTCC 2341  
QY 2342 ATCTATAACAAATCAATTAATTAATGACCCGAGGAAATTTTAAGGACAGAAAGT 2401  
DB 2342 ATCTATAACAAATCAATTAATTAATGACCCGAGGAAATTTTAAGGACAGAAAGT 2401  
QY 2402 AGCTTTGTTCAAGATTTTGTGCAATGGGAGCACTGTGATCATGAGAGACATCTGT 2461  
DB 2402 AGCTTTGTTCAAGATTTTGTGCAATGGGAGCACTGTGATCATGAGAGACATCTGT 2461  
QY 2462 TAGTGAGAGACCAAAACCTGTGTACCGTTTTCATGTATGAATTTTGTGTTAGT 2521  
DB 2462 TAGTGAGAGACCAAAACCTGTGTACCGTTTTCATGTATGAATTTTGTGTTAGT 2521  
QY 2522 TGCTTTAGCTAGCTGTGAGAGTCTGTGCTTTCTTAAGTGGGTATGAAAGGAGACATC 2581  
DB 2522 TGCTTTAGCTAGCTGTGAGAGTCTGTGCTTTCTTAAGTGGGTATGAAAGGAGACATC 2581  
QY 2582 TAACAAAATCATTAGATTAACAGCTCTCATGAGAGGAAACTAATCTCAATGTT 2641  
DB 2582 TAACAAAATCATTAGATTAACAGCTCTCATGAGAGGAAACTAATCTCAATGTT 2641  
QY 2642 TTAAGTAAATAAACTGTACTGCGCAAGTACTTTGAGCATTAATAAAAAAAAAAAG 2701  
DB 2642 TTAAGTAAATAAACTGTACTGCGCAAGTACTTTGAGCATTAATAAAAAAAAAAAG 2701  
QY 2702 GCGGCGCCG 2710  
DB 2702 GCGGCGCCG 2710

## RESULT 2

US-08-480-070C-8  
Sequence 8, Application US/08480070C  
Patent No. 6066498  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edwards LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,070C  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:

LENGTH: 2710 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 40..885  
US-08-480-070C-8  
Query Match 99.9%; Score 2706.6; DB 3; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 GTGACCCAGCGTCGCGGATTTCCCTCCCAAGTACATCTTTCAGTCTTACCCCTCA 61  
DB 2 GTGACCCAGCGTCGCGGATTTCCCTCCCAAGTACATCTTTCAGTCTTACCCCTCA 61  
QY 62 ACTGTGTCCTGCTGCTGCACTACTTTCGAAGGTCAATGGAGATGTTAAG 121  
DB 62 ACTGTGTCCTGCTGCTGCACTACTTTCGAAGGTCAATGGAGATGTTAAG 121  
QY 122 TTGAGTTGTGTAATAATGCTATCTGCTGCACTTCACTTCACTTCACTGAGAC 181  
DB 122 TTGAGTTGTGTAATAATGCTATCTGCTGCACTTCACTTCACTTCACTGAGAC 181  
QY 182 TTGAGCTTATGCTGTGGGCAAGGATTTCTCTTGTGTCAGTGTACATAGTTC 241  
DB 182 TTGAGCTTATGCTGTGGGCAAGGATTTCTCTTGTGTCAGTGTACATAGTTC 241  
QY 242 TCAGAACTGATGAAGAATAATGTGACATATCAGAAATCAGAGATPACAGCTTAAAG 301  
DB 242 TCAGAACTGATGAAGAATAATGTGACATATCAGAAATCAGAGATPACAGCTTAAAG 301  
QY 302 ATCTCAACAAAGAGATGTGTCTGTATCATTAAGAAATGACTCTGATGACCAATGGA 361  
DB 302 ATCTCAACAAAGAGATGTGTCTGTATCATTAAGAAATGACTCTGATGACCAATGGA 361  
QY 362 CCTACTGTCGAGATACAGTTCCTGCTTATTAATGATTAATAAATTAGACTGAAT 421  
DB 362 CCTACTGTCGAGATACAGTTCCTGCTTATTAATGATTAATAAATTAGACTGAAT 421  
QY 422 TAGACATCAAGACGCAAGGTCACTGCTGAGCTGAGCTGAGGAGCTCTACTACAG 481  
DB 422 TAGACATCAAGACGCAAGGTCACTGCTGAGCTGAGCTGAGGAGCTCTACTACAG 481  
QY 482 CTCTCCAAAGAACCTTAACACGAGAGAAATGTTTCAAGACACACACTGTGTACC 541  
DB 482 CTCTCCAAAGAACCTTAACACGAGAGAAATGTTTCAAGACACACACTGTGTACC 541  
QY 542 TCCATTAATTAACATGAAACAAATTTTCCATGAGGCTGATGAATTAAGACTCTGAG 601  
DB 542 TCCATTAATTAACATGAAACAAATTTTCCATGAGGCTGATGAATTAAGACTCTGAG 601  
QY 602 AAACGATCAGAACTGCTATCCAGATGGAAGGAGTCTGCTGGGTGACCTGGGAC 661  
DB 602 AAACGATCAGAACTGCTATCCAGATGGAAGGAGTCTGCTGGGTGACCTGGGAC 661  
QY 662 TTATCATTTGCTGCTTAATCTTAATGTTATTCCTGTAGAAAAGAAAGTTATCGAAT 721  
DB 662 TTATCATTTGCTGCTTAATCTTAATGTTATTCCTGTAGAAAAGAAAGTTATCGAAT 721  
QY 722 TGAGCTTTATTAACATGCGCAACTGCTCCAGAGAGGTTTGCAAAATGACAGAGAC 781  
DB 722 TGAGCTTTATTAACATGCGCAACTGCTCCAGAGAGGTTTGCAAAATGACAGAGAC 781  
QY 782 GGATTCGCTGAGGAAATATCTACACATCGAGAGAAAGTATATGAGTGAGAAAT 841  
DB 782 GGATTCGCTGAGGAAATATCTACACATCGAGAGAAAGTATATGAGTGAGAAAT 841  
QY 842 CAAATGAGTACTACTGCTAGCTCAACAGCCAGAGCAATCTGACCGGCTCTGAGCTCC 901  
DB 842 CAAATGAGTACTACTGCTAGCTCAACAGCCAGAGCAATCTGACCGGCTCTGAGCTCC 901

QY 902 ACTTTAAAGGCTGGCTTCATTTCTGACCTTGGATTTTCCCTTXXGGAATCTATG 961  
 DB 902 ACTTTAAAGGCTGGCTTCATTTCTGACCTTGGATTTTCCCTTXXGGAATCTATG 961  
 QY 962 ATATGCACTTGGCAACTCATTTGAGGTTCTGACCAAGCACTGAGAAAAGTTCC 1021  
 DB 962 ATATGCACTTGGCAACTCATTTGAGGTTCTGACCAAGCACTGAGAAAAGTTCC 1021  
 QY 1022 GTTTTCTGGGGATTAATTAATCAACAAGGGATTCGACTGATCATCTGATTAAT 1081  
 DB 1022 GTTTTCTGGGGATTAATTAATCAACAAGGGATTCGACTGATCATCTGATTAAT 1081  
 QY 1082 GCTCCATTTTATCCCTGAGTTTCAAGGATTCGATCCCACTGAGAACTTCAATCAG 1141  
 DB 1082 GCTCCATTTTATCCCTGAGTTTCAAGGATTCGATCCCACTGAGAACTTCAATCAG 1141  
 QY 1142 CGTGTGAAGCTCACTGCTGCTTCAATACATTAAGATGATGATGATCTTTGAGA 1201  
 DB 1142 CGTGTGAAGCTCACTGCTGCTTCAATACATTAAGATGATGATGATCTTTGAGA 1201  
 QY 1202 CATAGAGTTTGTGTATATCCCAAGCTCTGAAAGGTAGGGGAAATAAAGGCTAA 1261  
 DB 1202 CATAGAGTTTGTGTATATCCCAAGCTCTGAAAGGTAGGGGAAATAAAGGCTAA 1261  
 QY 1262 GATAGGAAGTGGGCTTGTGTATGATGTTGGAATCTTAAGAAGTGTGATCTTTCT 1321  
 DB 1262 GATAGGAAGTGGGCTTGTGTATGATGTTGGAATCTTAAGAAGTGTGATCTTTCT 1321  
 QY 1322 AGAGATTTCTGACCTTGAAAGATTAAGAAAAGCCAGGTGCAATGCTTAACAGATAT 1381  
 DB 1322 AGAGATTTCTGACCTTGAAAGATTAAGAAAAGCCAGGTGCAATGCTTAACAGATAT 1381  
 QY 1382 AACTTGGGAACCTTAGGCAAGAGGTGATTAAGTTCAAGGTCAAGGCTATGCTGTA 1441  
 DB 1382 AACTTGGGAACCTTAGGCAAGAGGTGATTAAGTTCAAGGTCAAGGCTATGCTGTA 1441  
 QY 1442 AGACTCTTCAMATCCAAAGAGAAATTAATGATGACACAGAGAGCTGAGATGA 1501  
 DB 1442 AGACTCTTCAMATCCAAAGAGAAATTAATGATGACACAGAGAGCTGAGATGA 1501  
 QY 1502 GGCTCGGACAGTGGGCTGATGTTGTAACAAGCAAGGATCTATATTTGATGCTAGACC 1561  
 DB 1502 GGCTCGGACAGTGGGCTGATGTTGTAACAAGCAAGGATCTATATTTGATGCTAGACC 1561  
 QY 1562 CCACATGAAAAGCTTAGGCTGTAAGCATGCTTGTAGACTCAAGAGATGAGAGTAA 1621  
 DB 1562 CCACATGAAAAGCTTAGGCTGTAAGCATGCTTGTAGACTCAAGAGATGAGAGTAA 1621  
 QY 1622 AGGCAACAAGATCCCGGGGCTTGGTGAGTCACTTACCTAGTGTGATGCCA 1681  
 DB 1622 AGGCAACAAGATCCCGGGGCTTGGTGAGTCACTTACCTAGTGTGATGCCA 1681  
 QY 1682 GTTCCAAAGATCCCTGTCTCAMAATGATGATGATGATGATGATGATGATGATG 1741  
 DB 1682 GTTCCAAAGATCCCTGTCTCAMAATGATGATGATGATGATGATGATGATGATG 1741  
 QY 1742 GGTGTCCTCTCTCTCAGAAAGATGACATGATGATGATGATGATGATGATGATG 1801  
 DB 1742 GGTGTCCTCTCTCTCAGAAAGATGACATGATGATGATGATGATGATGATGATG 1801  
 QY 1802 AC 1861  
 DB 1802 AC 1861  
 QY 1862 TCTCTAATCATGATCTCTACAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
 DB 1862 TCTCTAATCATGATCTCTACAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
 QY 1922 CATGCGAGAGCTGATGATTAATTTTCAAGCACTGAGAAAGCTGAGCAAGGCTGGA 1981  
 DB 1922 CATGCGAGAGCTGATGATTAATTTTCAAGCACTGAGAAAGCTGAGCAAGGCTGGA 1981

QY 1982 GAGTTCAGGAGCACTGTGTCCCAACATGCGACACTCTTCTTACACAGAAAAGTTACC 2041  
 DB 1982 GAGTTCAGGAGCACTGTGTCCCAACATGCGACACTCTTCTTACACAGAAAAGTTACC 2041  
 QY 2042 CCAGAGCACTGTGTCTGTGTAAAAGAAACCCTGCAAGGCAAACTTTGACTGTGTG 2101  
 DB 2042 CCAGAGCACTGTGTCTGTGTAAAAGAAACCCTGCAAGGCAAACTTTGACTGTGTG 2101  
 QY 2102 TGTCTAAGGGAACTGACTCAGACAACTTCTCATTTCTGAGAAACTGAGCTGTTTC 2161  
 DB 2102 TGTCTAAGGGAACTGACTCAGACAACTTCTCATTTCTGAGAAACTGAGCTGTTTC 2161  
 QY 2162 TGACAGAAACAAACCGGTGACTGAGACATAGAAAGGAGACTCTTGAGCAATCTATA 2221  
 DB 2162 TGACAGAAACAAACCGGTGACTGAGACATAGAAAGGAGACTCTTGAGCAATCTATA 2221  
 QY 2222 TAGTCAGCAAAATATTTCTTGGGAGACAGTGTGACCAAAATTTGCAAGCGGTGG 2281  
 DB 2222 TAGTCAGCAAAATATTTCTTGGGAGACAGTGTGACCAAAATTTGCAAGCGGTGG 2281  
 QY 2282 ACCTCAGTTTCACTGTGCTTACAGCTGCTGCCAGTGTGATCTGTGTGCTGCC 2341  
 DB 2282 ACCTCAGTTTCACTGTGCTTACAGCTGCTGCCAGTGTGATCTGTGTGCTGCC 2341  
 QY 2342 ATCTAATACAGATCAATTAATTAATGACCCGAGTGAATTAATGAGAGAAAGGT 2401  
 DB 2342 ATCTAATACAGATCAATTAATTAATGACCCGAGTGAATTAATGAGAGAAAGGT 2401  
 QY 2402 AGCTTGTTCAAAGATTTTCTTGGAGACAACTGTGTACATCAGAGACATCTGT 2461  
 DB 2402 AGCTTGTTCAAAGATTTTCTTGGAGACAACTGTGTACATCAGAGACATCTGT 2461  
 QY 2462 TAGTGAGACACCAAAACCTGTGTACAGCTTTTTCATGTATGAAATTTGTTAGGT 2521  
 DB 2462 TAGTGAGACACCAAAACCTGTGTACAGCTTTTTCATGTATGAAATTTGTTAGGT 2521  
 QY 2522 TGCTTCTAGCTAGCTGTGAGGTCCTGCTTCTTAGGTGATGAAAGGAGACATC 2581  
 DB 2522 TGCTTCTAGCTAGCTGTGAGGTCCTGCTTCTTAGGTGATGAAAGGAGACATC 2581  
 QY 2582 TAACAAAATCATTAAGATTAACAGCTCTCATGCAAGAGGAAATTAATCTCAATGTT 2641  
 DB 2582 TAACAAAATCATTAAGATTAACAGCTCTCATGCAAGAGGAAATTAATCTCAATGTT 2641  
 QY 2642 TTAAGTAAATTAATCTGCTGCAAGATCTTTGACGCTAATAAAAAAAAAAAG 2701  
 DB 2642 TTAAGTAAATTAATCTGCTGCAAGATCTTTGACGCTAATAAAAAAAAAAAG 2701  
 QY 2702 GGGGGGGGG 2710  
 DB 2702 GGGGGGGGG 2710

RESULT 3  
 US-08-829-525-8  
 ; Sequence 8, Application US/08829525  
 ; Patent No. 6084083  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Levinson, Douglas A.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 ; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
 ; NUMBER OF SEQUENCES: 38  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036/2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/829,525  
FILING DATE: 28-MAR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-081  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2710 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 40..885  
US-08-829-525-8

Query Match 99.9%; Score 2706.6; DB 3; Length 2710;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTCGACCCAGCGTCGCGATTTCCCTCCCAAGTACTGATTTTCAGGCTTACCCCTCA 61  
DB 2 GTCGACCCAGCGTCGCGATTTCCCTCCCAAGTACTGATTTTCAGGCTTACCCCTCA 61  
QY 62 ACTGTCCTGCTGCTGCTGCACTACTTTCGCAAGTCAATTTGAAGTGTAAAG 121  
DB 62 ACTGTCCTGCTGCTGCTGCACTACTTTCGCAAGTCAATTTGAAGTGTAAAG 121  
QY 122 TTGAGGTTGTTAAATGCTATCTGCTGCACTACTTTCGCAAGTCAATTTGAAG 181  
DB 122 TTGAGGTTGTTAAATGCTATCTGCTGCACTACTTTCGCAAGTCAATTTGAAG 181  
QY 182 TTGTCCTATGTCGTGGGCAAGGATTTCTGCTTGTGCAAGTCAATTTGAAGTGT 241  
DB 182 TTGTCCTATGTCGTGGGCAAGGATTTCTGCTTGTGCAAGTCAATTTGAAGTGT 241  
QY 242 TCAGACTGATGAAGAATGTGACATATCAGAAATTCAGAGATTCAGGCTAAAGGCG 301  
DB 242 TCAGACTGATGAAGAATGTGACATATCAGAAATTCAGAGATTCAGGCTAAAGGCG 301  
QY 302 ATCTGAAGAGAGATGTCCTGATCATATAAAGATGATCTTGTGATGCAATGGGA 361  
DB 302 ATCTGAAGAGAGATGTCCTGATCATATAAAGATGATCTTGTGATGCAATGGGA 361  
QY 362 CCTACTGTCGAGATAGTTCCTGCTTATGATGATTAATAAATTTAGAACTGAAT 421  
DB 362 CCTACTGTCGAGATAGTTCCTGCTTATGATGATTAATAAATTTAGAACTGAAT 421  
QY 422 TAGACATCAAGCAGCCCAAGTCTCAAGTCAAGTCCAGTGGGAGCTTACTAAG 481  
DB 422 TAGACATCAAGCAGCCCAAGTCTCAAGTCAAGTCCAGTGGGAGCTTACTAAG 481  
QY 482 CTTTCGAAGAGCCTTAACCAAGAGAGAAATGTTGAGAGACAGACACTGGTGACCC 541  
DB 482 CTTTCGAAGAGCCTTAACCAAGAGAGAAATGTTGAGAGACAGACACTGGTGACCC 541  
QY 542 TCCATTAATTAACATGAAGCAAAAAATTTCCACATGGGCTGATGAATTAAGACTGGAG 601

DB 542 TCCATTAATTAACATGAAGCAAAAAATTTCCACATGGGCTGATGAATTAAGACTGGAG 601  
QY 602 AAAGATCAAGACCTGATTCACATTTGAGAGGAGTCTGCTGGTTGACCCCTGGAC 661  
DB 602 AAAGATCAAGACCTGATTCACATTTGAGAGGAGTCTGCTGGTTGACCCCTGGAC 661  
QY 662 TTATCATGTCGTCTTAATCCTTAATGATTTTCCTGTAAGAAAGAAAGTATGAGTT 721  
DB 662 TTATCATGTCGTCTTAATCCTTAATGATTTTCCTGTAAGAAAGAAAGTATGAGTT 721  
QY 722 TGAGCTTATTAACATGAGCAATTTGCTCCAGAGAGGTTGGCAATGACAGAGCTCA 781  
DB 722 TGAGCTTATTAACATGAGCAATTTGCTCCAGAGAGGTTGGCAATGACAGAGCTCA 781  
QY 782 GGATTCGCTGAGAGAAATTTCTAACACATCGAGAGAAAGCTATTAAGTGGAGATT 841  
DB 782 GGATTCGCTGAGAGAAATTTCTAACACATCGAGAGAAAGCTATTAAGTGGAGATT 841  
QY 842 CAAATGAGTACTACTGCTAGCTCAAGCAGCCAGCCATCTGACCCGCTCTGAGCTGC 901  
DB 842 CAAATGAGTACTACTGCTAGCTCAAGCAGCCAGCCATCTGACCCGCTCTGAGCTGC 901  
QY 902 ACTTTAAAGGCTGCTTCAATTTCTGACTTTGATTTCCCTTTTGGAAAACTATGTG 961  
DB 902 ACTTTAAAGGCTGCTTCAATTTCTGACTTTGATTTCCCTTTTGGAAAACTATGTG 961  
QY 962 ATATGTCATCTGGCAACTCATTTGAGAGTCTGACACAGCCACTGAGAAAAAGATTCCA 1021  
DB 962 ATATGTCATCTGGCAACTCATTTGAGAGTCTGACACAGCCACTGAGAAAAAGATTCCA 1021  
QY 1022 GTTTTCGAGGATTAATTAATCAACAGGAGGATTTGAGTACTGATCTATGATTAAT 1081  
DB 1022 GTTTTCGAGGATTAATTAATCAACAGGAGGATTTGAGTACTGATCTATGATTAAT 1081  
QY 1082 GCTCCATTTTATCCCTGAGTTTCAAGGATTCGATCTCCACTCCAGACTTCAATCATG 1141  
DB 1082 GCTCCATTTTATCCCTGAGTTTCAAGGATTCGATCTCCACTCCAGACTTCAATCATG 1141  
QY 1142 CGTGTGAAGTCACTGCTGCTTTCATATCATTAGAAATGTTAGTGTATGCTTTGGA 1201  
DB 1142 CGTGTGAAGTCACTGCTGCTTTCATATCATTAGAAATGTTAGTGTATGCTTTGGA 1201  
QY 1202 CATAGAGTTGTTGATATCCGAAAGCTCTGTAACAGTAGGGGAAATAAGGGCTTA 1261  
DB 1202 CATAGAGTTGTTGATATCCGAAAGCTCTGTAACAGTAGGGGAAATAAGGGCTTA 1261  
QY 1262 GATAGAGGTCGCTTTTGTGATGTTGAAAAATCTTAAGAAGTGTAGCTTTTCT 1321  
DB 1262 GATAGAGGTCGCTTTTGTGATGTTGAAAAATCTTAAGAAGTGTAGCTTTTCT 1321  
QY 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGATGCTTAAACAGATAT 1381  
DB 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGATGCTTAAACAGATAT 1381  
QY 1382 AACTTGGGAACTTAAGAGAGAGGTTAAGTTCAAGSTAGCCAGGGCTATAGCTGTA 1441  
DB 1382 AACTTGGGAACTTAAGAGAGAGGTTAAGTTCAAGSTAGCCAGGGCTATAGCTGTA 1441  
QY 1442 AGACTGTCATCAATCCAAAGCAAGAAATAAACAATAGAGACAGAGAGGCTGAGATGA 1501  
DB 1442 AGACTGTCATCAATCCAAAGCAAGAAATAAACAATAGAGACAGAGAGGCTGAGATGA 1501  
QY 1502 GGCCTGGAAGTGAAGTGCATTTGTTGTAACACAGAGAACTTAATTTGATTCGTAAGC 1561  
DB 1502 GGCCTGGAAGTGAAGTGCATTTGTTGTAACACAGAGAACTTAATTTGATTCGTAAGC 1561  
QY 1562 CCACATGAAGAAAGCTAGGCTGTTGAGACATGCTTTGATCTCAAGAGATGAGAGGTA 1621  
DB 1562 CCACATGAAGAAAGCTAGGCTGTTGAGACATGCTTTGATCTCAAGAGATGAGAGGTA 1621  
QY 1622 AGGCAACAAGATCCCGGGGCTTGCCTGAGTCAAGCTTAACTGATGCTGAGTTCCA 1681





Db 242 TCAGACATGATGAAAGAAATGTGACATATCAGAAATCCAGATACCAAGCTAAAGGCG 301  
 Qy 302 ATCTCAAGAAAGGAGATGTGTCTGATCATATAAAGATGACTCTGATGATCCATGGGA 361  
 Db 302 ATCTCAAGAAAGGAGATGTGTCTGATCATATAAAGATGACTCTGATGATCCATGGGA 361  
 Qy 362 CCTACTGTGACGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421  
 Db 362 CCTACTGTGACGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421  
 Qy 422 TAGACATCAAGACCCCAAGGCTCACTCCAGCTCAAGCTCCAGCTCCAGCTCCAGCT 481  
 Db 422 TAGACATCAAGACCCCAAGGCTCACTCCAGCTCAAGCTCCAGCTCCAGCTCCAGCT 481  
 Qy 482 CTCTCCCAAGAACCTTAACCAAGGAGAGAAATGTTTCAAGACACAGACACTGTGACCC 541  
 Db 482 CTCTCCCAAGAACCTTAACCAAGGAGAGAAATGTTTCAAGACACAGACACTGTGACCC 541  
 Qy 542 TCATATATTAACATGGAACAAATTTTCCACATGGGCTGATGATTAAGAACTCTGGAG 601  
 Db 542 TCATATATTAACATGGAACAAATTTTCCACATGGGCTGATGATTAAGAACTCTGGAG 601  
 Qy 602 AAACGATAGAACTGCTATCCACATTTGAGAGTGTCTCTGCTGGTTGACCTTGGAC 661  
 Db 602 AAACGATAGAACTGCTATCCACATTTGAGAGTGTCTCTGCTGGTTGACCTTGGAC 661  
 Qy 662 TTAATATGTTGTCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTA 721  
 Db 662 TTAATATGTTGTCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTA 721  
 Qy 722 TGAGCTTATTAACATGGAACAAATTTTCCACATGGGCTGATGATTAAGAACTCTGGAG 781  
 Db 722 TGAGCTTATTAACATGGAACAAATTTTCCACATGGGCTGATGATTAAGAACTCTGGAG 781  
 Qy 782 GATATGCTCTGAGAGAAATATCTAACCATGAGAGAAACGATATATGAGTGAAT 841  
 Db 782 GATATGCTCTGAGAGAAATATCTAACCATGAGAGAAACGATATATGAGTGAAT 841  
 Qy 842 CAAAGATTAACATGGAACAAATTTTCCACATGGGCTGATGATTAAGAACTCTGGAG 901  
 Db 842 CAAAGATTAACATGGAACAAATTTTCCACATGGGCTGATGATTAAGAACTCTGGAG 901  
 Qy 902 ACTTTTAAAGGCTGCTCTTCAATTTTGAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 961  
 Db 902 ACTTTTAAAGGCTGCTCTTCAATTTTGAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 961  
 Qy 962 ATATGCTCACTTGGCAACCTCATTTGAGAGTGTCTGACACAGCCACTGAGAAAGATTCA 1021  
 Db 962 ATATGCTCACTTGGCAACCTCATTTGAGAGTGTCTGACACAGCCACTGAGAAAGATTCA 1021  
 Qy 1022 GTTTTCTGGGAGTAATTAATCAACAAAGGAGATTGCACTGTATCACTGATGATTAAG 1081  
 Db 1022 GTTTTCTGGGAGTAATTAATCAACAAAGGAGATTGCACTGTATCACTGATGATTAAG 1081  
 Qy 1082 GCTCCATTTTATCCCTGAGTTTCAAGGATCGGATCTCCACTCCAGAGCTTCAATCATG 1141  
 Db 1082 GCTCCATTTTATCCCTGAGTTTCAAGGATCGGATCTCCACTCCAGAGCTTCAATCATG 1141  
 Qy 1142 CGTGTGGAAGCTCACTGCTGCTTTCAATCAATTAAGATGTTAGTGTGTTTGAAGA 1201  
 Db 1142 CGTGTGGAAGCTCACTGCTGCTTTCAATCAATTAAGATGTTAGTGTGTTTGAAGA 1201  
 Qy 1202 CATTAAGGTTTGTGTATATCCGCAAGGCTCTGTAACAGTTAGGGGAGATTAAGGGCTAA 1261  
 Db 1202 CATTAAGGTTTGTGTATATCCGCAAGGCTCTGTAACAGTTAGGGGAGATTAAGGGCTAA 1261  
 Qy 1262 GATTAAGGTTTGTGTATATCCGCAAGGCTCTGTAACAGTTAGGGGAGATTAAGGGCTAA 1321  
 Db 1262 GATTAAGGTTTGTGTATATCCGCAAGGCTCTGTAACAGTTAGGGGAGATTAAGGGCTAA 1321  
 Qy 1322 AGAGATTTCTGACCTTGAAGATTAAAGAAAGCCAGGTGCAATGCTTAACACGATAT 1381  
 Db 1322 AGAGATTTCTGACCTTGAAGATTAAAGAAAGCCAGGTGCAATGCTTAACACGATAT 1381

Db 1382 AGAGATTTCTGACCTTGAAGATTAAAGAAAGCCAGGTGCAATGCTTAACACGATAT 1381  
 Qy 1382 AACTTGGGAACCTTAAAGCAGAGGGTGAATGTTCAAGCTCAGCCAGGGCTATGCTGTA 1441  
 Db 1382 AACTTGGGAACCTTAAAGCAGAGGGTGAATGTTCAAGCTCAGCCAGGGCTATGCTGTA 1441  
 Qy 1442 AGACTGTCTCAGATCCAAAGACGAAATTAACATATAGACAGCAGAGGCTGAGATGA 1501  
 Db 1442 AGACTGTCTCAGATCCAAAGACGAAATTAACATATAGACAGCAGAGGCTGAGATGA 1501  
 Qy 1502 GAGTGGAGCAGTGAAGTGCATTTGTATACAGACAGAGAACTATATTTGATTCGTAGAC 1561  
 Db 1502 GAGTGGAGCAGTGAAGTGCATTTGTATACAGACAGAGAACTATATTTGATTCGTAGAC 1561  
 Qy 1562 CCACATGAAAGAGCTAGGCTGTATAGACAGCTTGTATGATCTCAAGAGATGAGAGTAA 1621  
 Db 1562 CCACATGAAAGAGCTAGGCTGTATAGACAGCTTGTATGATCTCAAGAGATGAGAGTAA 1621  
 Qy 1622 AGGCAACAAGATCCCGGGGCTTGCCTGACAGTCAAGCTTACCTAGGTGATGCTCAA 1681  
 Db 1622 AGGCAACAAGATCCCGGGGCTTGCCTGACAGTCAAGCTTACCTAGGTGATGCTCAA 1681  
 Qy 1682 GTTCAACAAGATCCCTGTCTCAGAGTAAATGATGATGATGATGATGATGATGATGATGAT 1741  
 Db 1682 GTTCAACAAGATCCCTGTCTCAGAGTAAATGATGATGATGATGATGATGATGATGATGAT 1741  
 Qy 1742 GGTGTCT 1801  
 Db 1742 GGTGTCT 1801  
 Qy 1802 ACAACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGATG 1861  
 Db 1802 ACAACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGATG 1861  
 Qy 1862 TCTCTAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1921  
 Db 1862 TCTCTAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1921  
 Qy 1922 CATGAG 1981  
 Db 1922 CATGAG 1981  
 Qy 1982 GAGTTCAAGAGACACTGTCCTCCCAACACTGACAGACTCTTCTTACACAAAGAAAGTTACC 2041  
 Db 1982 GAGTTCAAGAGACACTGTCCTCCCAACACTGACAGACTCTTCTTACACAAAGAAAGTTACC 2041  
 Qy 2042 CGCAAGCAGCTGCTGTCTGTAAAGAAAGCCCTGCGAAAGGCAACTTTGACTGTGTG 2101  
 Db 2042 CGCAAGCAGCTGCTGTCTGTAAAGAAAGCCCTGCGAAAGGCAACTTTGACTGTGTG 2101  
 Qy 2102 TGCTCAAGGAGACATGATCTCAGACAACTTCTCTCATTTCTGAGAGAAACTGAGAGCTGTTT 2161  
 Db 2102 TGCTCAAGGAGACATGATCTCAGACAACTTCTCTCATTTCTGAGAGAAACTGAGAGCTGTTT 2161  
 Qy 2162 TGAAGAGAGACAAACCGGTGACTGGGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2221  
 Db 2162 TGAAGAGAGACAAACCGGTGACTGGGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2221  
 Qy 2222 TAGTCAGCAAAATATTTCTTGGAGAGACAGTGTGTCACAAATTTGATTTTCAAGCCGGTGG 2281  
 Db 2222 TAGTCAGCAAAATATTTCTTGGAGAGACAGTGTGTCACAAATTTGATTTTCAAGCCGGTGG 2281  
 Qy 2282 AACTCAAGTTTCAATCTGAGTTTCAAGTGTGCTGCTCCAGTCCCTTGAATCTGTGTGCTGCC 2341  
 Db 2282 AACTCAAGTTTCAATCTGAGTTTCAAGTGTGCTGCTCCAGTCCCTTGAATCTGTGTGCTGCC 2341  
 Qy 2342 ATCTATAACAGATCAATTAATTAAGACCCGAGTGAATATTAAGTGAAGCAAGAGT 2401  
 Db 2342 ATCTATAACAGATCAATTAATTAAGACCCGAGTGAATATTAAGTGAAGCAAGAGT 2401  
 Qy 2402 AGCTTTGTTCAAGATTTTCTTGAATGGAGACACTGTGTACATCAGAGACATCTGT 2461  
 Db 2402 AGCTTTGTTCAAGATTTTCTTGAATGGAGACACTGTGTACATCAGAGACATCTGT 2461

QY 2462 TAGTGAAGACACCAAAACCTGTGTACCCGTTTTCATGTATGAATTTTGTGTAGGT 2521  
Db 2462 TAGTGAAGACACCAAAACCTGTGTACCCGTTTTCATGTATGAATTTTGTGTAGGT 2521  
QY 2522 TGCTTACCTAGCTGTGTAGAGTCTGTGGCTTTCTTAGGTGTATGAAGAGGAGACATC 2581  
Db 2522 TGCTTACCTAGCTGTGTAGAGTCTGTGGCTTTCTTAGGTGTATGAAGAGGAGACATC 2581  
QY 2582 TAACAAAATTCATTAGATTAACAGTCTTCATGACAGAGGAAACTAATCTCAATGTT 2641  
Db 2582 TAACAAAATTCATTAGATTAACAGTCTTCATGACAGAGGAAACTAATCTCAATGTT 2641  
QY 2642 TTTAAATTAATAAAGTGTACTGTGCAAAAGTCTTGTGACATTAATAAAAAAAAAAAG 2701  
Db 2642 TTTAAATTAATAAAGTGTACTGTGCAAAAGTCTTGTGACATTAATAAAAAAAAAAAG 2701  
QY 2702 GGGGGCCGC 2710  
Db 2702 GGGGGCCGC 2710

RESULT 5  
US-08-937-399-8  
Sequence 8, Application US/08937399  
Patent No. 6288218  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/937,399  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Cornuzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2710 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 40..885  
US-08-937-399-8

Query Match 99.9%; Score 2706.6; DB 3; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTCGACCCACGGGCTCCGAGTTTCCCTCCCAAGTACTATGTTTTCAGGCTTACCTCA 61  
Db 2 GTCGACCCACGGGCTCCGAGTTTCCCTCCCAAGTACTATGTTTTCAGGCTTACCTCA 61  
QY 62 ACTGTGCTGTGCTGTGCAACTACTATTGCAAGGTCAATGGAAGTGTATTAAG 121  
Db 62 ACTGTGCTGTGCTGTGCAACTACTATTGCAAGGTCAATGGAAGTGTATTAAG 121  
QY 122 TTGAGGTGTATTAATAAGTCTATTCGCTCGAGTTACCTTACTTATCTGGACAC 181  
Db 122 TTGAGGTGTATTAATAAGTCTATTCGCTCGAGTTACCTTACTTATCTGGACAC 181  
QY 182 TTGTGCTATGTGTGTGGGCAAGGATTCGTCTTGTGTCAGTGTACAGTATCAATGAGTTGC 241  
Db 182 TTGTGCTATGTGTGTGGGCAAGGATTCGTCTTGTGTCAGTGTACAGTATCAATGAGTTGC 241  
QY 242 TCAGAACTGTATGAAGAAATGTGACATATCAGAAATCCAGAGATACCAAGGCGC 301  
Db 242 TCAGAACTGTATGAAGAAATGTGACATATCAGAAATCCAGAGATACCAAGGCGC 301  
QY 302 ATCTCAACAAAGAGATGTCTGTGATCATTAAGAAATGTGATCTGTGATGACATGGA 361  
Db 302 ATCTCAACAAAGAGATGTCTGTGATCATTAAGAAATGTGATCTGTGATGACATGGA 361  
QY 362 CCTACTGTGCAAGTATACAGTTCCTGTGCTTATGATATGATTAATAATAGAACTGAAT 421  
Db 362 CCTACTGTGCAAGTATACAGTTCCTGTGCTTATGATATGATTAATAATAGAACTGAAT 421  
QY 422 TAGAATCAAGACGCAAGGTCTCTGATCAATGCACTGCGGAGCTTACTACAG 481  
Db 422 TAGAATCAAGACGCAAGGTCTCTGATCAATGCACTGCGGAGCTTACTACAG 481  
QY 482 CTCTCCCAAGAACCTTAACCAAGGAGAGAAATGTTCAGAGACACAGACACTGTGACCC 541  
Db 482 CTCTCCCAAGAACCTTAACCAAGGAGAGAAATGTTCAGAGACACAGACACTGTGACCC 541  
QY 542 TCCATTAATTAACAGTGAACCAAAATTTCCACATGGGCTGTGATTAATAGACTGTGAG 601  
Db 542 TCCATTAATTAACAGTGAACCAAAATTTCCACATGGGCTGTGATTAATAGACTGTGAG 601  
QY 602 AAACGATCGAATCTGTATCAATTTGAGTGGAGTCTCTGCTGCTTGAACCTTGGCAC 661  
Db 602 AAACGATCGAATCTGTATCAATTTGAGTGGAGTCTCTGCTGCTTGAACCTTGGCAC 661  
QY 662 TTATCATTTGGTCTTAATCCTTAATGTAATTCCTGTGAAGAAAGAAAGTATGAGTT 721  
Db 662 TTATCATTTGGTCTTAATCCTTAATGTAATTCCTGTGAAGAAAGAAAGTATGAGTT 721  
QY 722 TGAGCTTTATACACTGAGCACTTGCTCCAGAGAGGTGTGCAATGAGAGGAGTGA 781  
Db 722 TGAGCTTTATACACTGAGCACTTGCTCCAGAGAGGTGTGCAATGAGAGGAGTGA 781  
QY 782 GGAATTCCTGTGAGAGAAATATCTACACCATGAGAGAAAGTATATGAGTGAAGTT 841  
Db 782 GGAATTCCTGTGAGAGAAATATCTACACCATGAGAGAAAGTATATGAGTGAAGTT 841  
QY 842 CAAATGAGATCACTGCTACGTCACACAGACAGACATCCGACCGCTCGAGCTGGC 901  
Db 842 CAAATGAGATCACTGCTACGTCACACAGACAGACATCCGACCGCTCGAGCTGGC 901  
QY 902 ACTTTTAAAGGCTCGCTCATTTCTGACTTGTATTTCCCTTTTGTGAAAACATATGTG 961  
Db 902 ACTTTTAAAGGCTCGCTCATTTCTGACTTGTATTTCCCTTTTGTGAAAACATATGTG 961  
QY 962 ATATGTCACTTGCAACCTCATTTGAGAGGTTTGACACACAGCACTGAGAAAAGATTCCA 1021  
Db 962 ATATGTCACTTGCAACCTCATTTGAGAGGTTTGACACACAGCACTGAGAAAAGATTCCA 1021

QY 1022 GTTTTCGAGGATTAATACTCAGAGGGGATTCAGCTGTAACTCATGTCTACATTGAAT 1081  
DB 1022 GTTTTCGAGGATTAATACTCAGAGGGGATTCAGCTGTAACTCATGTCTACATTGAAT 1081  
QY 1082 GCTTCATTTTATCCCTGAGTTTCAGGATTCGATCTCCCACTCCAGAGACTTCATCATG 1141  
DB 1082 GCTTCATTTTATCCCTGAGTTTCAGGATTCGATCTCCCACTCCAGAGACTTCATCATG 1141  
QY 1142 CGTGTGAAGCTCATCTGCTGCTTTCATCATTAAGGATTCGATCTCCCACTCCAGAG 1201  
DB 1142 CGTGTGAAGCTCATCTGCTGCTTTCATCATTAAGGATTCGATCTCCCACTCCAGAG 1201  
QY 1202 CATAGAGGTTGTGTATATCCGAAAGCTCTGAAACAGGTGAGGGAATTAAGGCTTAA 1261  
DB 1202 CATAGAGGTTGTGTATATCCGAAAGCTCTGAAACAGGTGAGGGAATTAAGGCTTAA 1261  
QY 1262 GATAGAGGTTGTGTATATCCGAAAGCTCTGAAACAGGTGAGGGAATTAAGGCTTAA 1321  
DB 1262 GATAGAGGTTGTGTATATCCGAAAGCTCTGAAACAGGTGAGGGAATTAAGGCTTAA 1321  
QY 1322 AGAGATTTCTGACCTGTAAGATTAAGAAAGCCAGGTCGATTCGATTCACAGATAT 1381  
DB 1322 AGAGATTTCTGACCTGTAAGATTAAGAAAGCCAGGTCGATTCGATTCACAGATAT 1381  
QY 1382 AACTTGGGAACTTATAGGAGGAGGATTAAGATTCAGAGGCTTATGCTGTGTA 1441  
DB 1382 AACTTGGGAACTTATAGGAGGAGGATTAAGATTCAGAGGCTTATGCTGTGTA 1441  
QY 1442 AGACTGTCTCAGCATTCAGAAAGCAAAATTAACATAGAGACAGAGAGGCTGAGATGA 1501  
DB 1442 AGACTGTCTCAGCATTCAGAAAGCAAAATTAACATAGAGACAGAGAGGCTGAGATGA 1501  
QY 1502 GGGCTGGACAGTGGAGTGTGTAAGAGACAGAGGATTCATATTTGATCTGAGAC 1561  
DB 1502 GGGCTGGACAGTGGAGTGTGTAAGAGACAGAGGATTCATATTTGATCTGAGAC 1561  
QY 1562 CCACATGAAAAAGCTAGAGCTGTAGAGCATCTGTAGACTCAAGAGATGAGAGATTA 1621  
DB 1562 CCACATGAAAAAGCTAGAGCTGTAGAGCATCTGTAGACTCAAGAGATGAGAGATTA 1621  
QY 1622 AGGCAACACAGATCCCCGGGGCTTGGCTGCACTCAGCTTACGCTTACGCTGCTCA 1681  
DB 1622 AGGCAACACAGATCCCCGGGGCTTGGCTGCACTCAGCTTACGCTTACGCTGCTCA 1681  
QY 1682 GTTCCAGAGATCCCTGCTCAGAGATGAGTGTGATCTGAGGCTGATCTCCATGAG 1741  
DB 1682 GTTCCAGAGATCCCTGCTCAGAGATGAGTGTGATCTGAGGCTGATCTCCATGAG 1741  
QY 1742 GGTTCCTCTCTCTCAGAGAGATGAGTGTGATCTGAGGCTGATCTCCATGAG 1801  
DB 1742 GGTTCCTCTCTCTCAGAGAGATGAGTGTGATCTGAGGCTGATCTCCATGAG 1801  
QY 1802 ACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGATG 1861  
DB 1802 ACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGATG 1861  
QY 1862 TCTCTATATCATGATCTCTACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
DB 1862 TCTCTATATCATGATCTCTACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
QY 1922 CATGCGAGAGAGTCCAGTAAATTTATTCAGAGCTCAGAGGCTGAGAGAGAGCTGGA 1981  
DB 1922 CATGCGAGAGAGTCCAGTAAATTTATTCAGAGCTCAGAGGCTGAGAGAGAGCTGGA 1981  
QY 1982 GAGTTCAAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2041  
DB 1982 GAGTTCAAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2041  
QY 2042 CGCAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2101  
DB 2042 CGCAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2101  
QY 2102 TGCTCAAGGGGAACTGAGCTCAGACAACTTCTCATCTCTGAGAGAACTGAGCTGTTTC 2161

DB 2102 TGCTCAAGGGGAACTGAGCTCAGACAACTTCTCATCTCTGAGAGAACTGAGCTGTTTC 2161  
QY 2162 TGACAGAGAGCAACCCGCTGAGCTGGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 2221  
DB 2162 TGACAGAGAGCAACCCGCTGAGCTGGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 2221  
QY 2222 TAGTCAGAGAGATATCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2281  
DB 2222 TAGTCAGAGAGATATCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2281  
QY 2282 AGCTGAGTTCATCTGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2341  
DB 2282 AGCTGAGTTCATCTGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2341  
QY 2342 ATCTATTAACAGATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2401  
DB 2342 ATCTATTAACAGATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2401  
QY 2402 AGCTTGTTCAGAGATTTTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2461  
DB 2402 AGCTTGTTCAGAGATTTTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2461  
QY 2462 TAGTGAAGACACAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2521  
DB 2462 TAGTGAAGACACAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2521  
QY 2522 TGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2581  
DB 2522 TGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2581  
QY 2582 TTAACAAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2641  
DB 2582 TTAACAAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2641  
QY 2642 TTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2701  
DB 2642 TTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2701  
QY 2702 GGGGGGGG 2710  
DB 2702 GGGGGGGG 2710

RESULT 6  
US-09-310-367-8  
Sequence 8, Application US/09310367  
Patent No. 6414117  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/310,367  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/829,525  
FILING DATE: 28-MAR-1997  
APPLICATION NUMBER: US 08/609,583





Dh 302 ATCTCAACAAAGAGATGTCTCTGATCAATAAGAAATGTGACTCTGATGACCAATGGCA 361  
Qy 362 CCTACTGCTGCAGAGATACAGTTCCCTGGTCTTATGATATGATTAATAAATTAGAACTGAAT 421  
Dh 362 CCTACTGCTGCAGAGATACAGTTCCCTGGTCTTATGATATGATTAATAAATTAGAACTGAAT 421  
Qy 422 TAGACATCAAAAGAGCCAAAGGTCACTCCAGCTCAGACTGCGCCATGGGGACTCTACACAG 481  
Dh 422 TAGACATCAAAAGAGCCAAAGGTCACTCCAGCTCAGACTGCGCCATGGGGACTCTACACAG 481  
Qy 482 CTCTCCAAAGAACCTTACACGAGAGAAATGGTTGAGAGACACAGACTGAGTGGAC 541  
Dh 482 CTCTCCAAAGAACCTTACACGAGAGAAATGGTTGAGAGACACAGACTGAGTGGAC 541  
Qy 542 TCCATATAACAAATGGAACAAAATTTCCATGAGGCTGATGAAATTAAGACTCTGGAG 601  
Dh 542 TCCATATAACAAATGGAACAAAATTTCCATGAGGCTGATGAAATTAAGACTCTGGAG 601  
Qy 602 AAACGATCAGAACTGCTATCCACATGGAGTGGAGTCTCTGCTGGGTGACCTGGCAC 661  
Dh 602 AAACGATCAGAACTGCTATCCACATGGAGTGGAGTCTCTGCTGGGTGACCTGGCAC 661  
Qy 662 TTAATCATGCTCTTATCTTAAATGATTCCTGTAGAGAAAAGAGTTATCGAGTT 721  
Dh 662 TTAATCATGCTCTTATCTTAAATGATTCCTGTAGAGAAAAGAGTTATCGAGTT 721  
Qy 722 TGAGCCTTATTAACATGAGCACTGGCAACTTGGCTCAGAGAGGGTTGGCAATGACAGACATCA 781  
Dh 722 TGAGCCTTATTAACATGAGCACTGGCAACTTGGCTCAGAGAGGGTTGGCAATGACAGACATCA 781  
Qy 782 GGAATGCGCTGAGAGAAATATCTACACCATCGAGAGAAAGATATGATGAGTGAAT 841  
Dh 782 GGAATGCGCTGAGAGAAATATCTACACCATCGAGAGAAAGATATGATGAGTGAAT 841  
Qy 842 CAATGAGTACTACTGCTACATGCTCAACAGCCAGCAGCATCTGACCGCTCTGACATGCC 901  
Dh 842 CAATGAGTACTACTGCTACATGCTCAACAGCCAGCAGCATCTGACCGCTCTGACATGCC 901  
Qy 902 ACTTTTAAAGGCTGCGCTTCAATTTCTGACTTTGGTATTTCCCTTTTGGAAAATCTATG 961  
Dh 902 ACTTTTAAAGGCTGCGCTTCAATTTCTGACTTTGGTATTTCCCTTTTGGAAAATCTATG 961  
Qy 962 ATATGTCATTTGGCACTCATTTGGAGTTCTGACCAAGCCACTGAGAAAAGAGTTCCA 1021  
Dh 962 ATATGTCATTTGGCACTCATTTGGAGTTCTGACCAAGCCACTGAGAAAAGAGTTCCA 1021  
Qy 1022 GTTTTCTGGGGATATTAATCACTCAAGAGGGATTCGACTGTAATCTCATCTGTAAT 1081  
Dh 1022 GTTTTCTGGGGATATTAATCACTCAAGAGGGATTCGACTGTAATCTCATCTGTAAT 1081  
Qy 1082 GTTCCATTTTATCCCTGAGTTTCAAGGATCGATCTCCACTCAGAGACTTCAATCAG 1141  
Dh 1082 GTTCCATTTTATCCCTGAGTTTCAAGGATCGATCTCCACTCAGAGACTTCAATCAG 1141  
Qy 1142 CGTGTGAACCTCACTGCTCTTCACTACATTAAGGAAATGTTAGTGTGATCTTTGAGA 1201  
Dh 1142 CGTGTGAACCTCACTGCTCTTCACTACATTAAGGAAATGTTAGTGTGATCTTTGAGA 1201  
Qy 1202 CATAGAGTTTGTGATATCCGCAAGGCTCTGAAACAGGTAGGGGGAATAAAGGGGTAA 1261  
Dh 1202 CATAGAGTTTGTGATATCCGCAAGGCTCTGAAACAGGTAGGGGGAATAAAGGGGTAA 1261  
Qy 1262 GATAGGAAGGTGCGGTCTTTGTGATGTTGAAAATCTTAAAGAGTTGATGCTTTTCT 1321  
Dh 1262 GATAGGAAGGTGCGGTCTTTGTGATGTTGAAAATCTTAAAGAGTTGATGCTTTTCT 1321  
Qy 1322 AGAGATTTCTGACTTGAATAAAGATTAAGAAAAAGCAGAGTGCATATCTTAACAGAT 1381  
Dh 1322 AGAGATTTCTGACTTGAATAAAGATTAAGAAAAAGCAGAGTGCATATCTTAACAGAT 1381  
Qy 1382 AACTTGGGAACCTTAGGCAAGAGGGTGAATAAGTTCAAGGTCAGCAGGGCTATGCTGTGA 1441  
Dh 1382 AACTTGGGAACCTTAGGCAAGAGGGTGAATAAGTTCAAGGTCAGCAGGGCTATGCTGTGA 1441

Qy 1442 AGACTGTCTCAGCATTCAAAAGAGAAAATTAACATAGAGACAGCAGAGAGCTGGAGATGA 1501  
Dh 1442 AGACTGTCTCAGCATTCAAAAGAGAAAATTAACATAGAGACAGCAGAGAGCTGGAGATGA 1501  
Qy 1502 GGTGGGAGAGAGGTGATGATGTAACAGAGCAGAGAAATCTATATTTGATGCTGAGCC 1561  
Dh 1502 GGTGGGAGAGAGGTGATGATGTAACAGAGCAGAGAAATCTATATTTGATGCTGAGCC 1561  
Qy 1562 CCACATGAAAAAGCTGAGCCTGGTAGAGCATGCTTGTAGACTCAAGAGATGAGAGTAA 1621  
Dh 1562 CCACATGAAAAAGCTGAGCCTGGTAGAGCATGCTTGTAGACTCAAGAGATGAGAGTAA 1621  
Qy 1622 AGGCAACAAGATCCCCGGGGGCTGGCTGCAAGTCAAGCTTAGCTAGTGTGCTCA 1681  
Dh 1622 AGGCAACAAGATCCCCGGGGGCTGGCTGCAAGTCAAGCTTAGCTAGTGTGCTCA 1681  
Qy 1682 GTTCCAAAGAGTCCCTGTCCAAAGTGAATGSRCTGAGTATCTGGCGCATGCTCATGGG 1741  
Dh 1682 GTTCCAAAGAGTCCCTGTCCAAAGTGAATGSRCTGAGTATCTGGCGCATGCTCATGGG 1741  
Qy 1742 GGTGTCTCTCTCTCAGAGAGACATGCAATGWCCTTGACACACACACACACAC 1801  
Dh 1742 GGTGTCTCTCTCTCAGAGAGACATGCAATGWCCTTGACACACACACACACAC 1801  
Qy 1802 ACACACACACACACACACACACACACATGATGATGAGTCTCTGCTGCTTACC 1861  
Dh 1802 ACACACACACACACACACACACACACATGATGATGAGTCTCTGCTGCTTACC 1861  
Qy 1862 TCTCTATAATGATATCTTACAGAGCTGCTGCTGCTCTGTGTAAGATGAGTGGAG 1921  
Dh 1862 TCTCTATAATGATATCTTACAGAGCTGCTGCTGCTCTGTGTAAGATGAGTGGAG 1921  
Qy 1922 CATGGCAGAGCAGTCCAGTATTTATTCAGACCTCAGAGAGCTGGAGCAGAGCGTGA 1981  
Dh 1922 CATGGCAGAGCAGTCCAGTATTTATTCAGACCTCAGAGAGCTGGAGCAGAGCGTGA 1981  
Qy 1982 GAGTTGAGAGCACTGTGCTCCCAACTGCTCAGACTTTCTTACACAGAAAAAGTTACC 2041  
Dh 1982 GAGTTGAGAGCACTGTGCTCCCAACTGCTCAGACTTTCTTACACAGAAAAAGTTACC 2041  
Qy 2042 CGCAAGCAGCTGCTGTGTAAAGAGAAACCTGGGAAAGCAATTTGACTGTTGTG 2101  
Dh 2042 CGCAAGCAGCTGCTGTGTAAAGAGAAACCTGGGAAAGCAATTTGACTGTTGTG 2101  
Qy 2102 TGCTCAAGGGGAACTGACTCAACAACTTCTCATTTCTGAGAGAACTGAGCTGTTTC 2161  
Dh 2102 TGCTCAAGGGGAACTGACTCAACAACTTCTCATTTCTGAGAGAACTGAGCTGTTTC 2161  
Qy 2162 TGACAGAAAGCAACCGGTGACTGGGACATAGAGAGCAGAGCTTTGACAGCAATCTAT 222  
Dh 2162 TGACAGAAAGCAACCGGTGACTGGGACATAGAGAGCAGAGCTTTGACAGCAATCTAT 222  
Qy 2222 TAGTCAGAAAATATTCTTTGGGAGAGCAGTGCCTCAACAAATTTGATTTCAAGCCGGTGG 2281  
Dh 2222 TAGTCAGAAAATATTCTTTGGGAGAGCAGTGCCTCAACAAATTTGATTTCAAGCCGGTGG 2281  
Qy 2282 ACCTCAGTTTCACTGCTTAAAGCTGCTGCCAGTGCCTTGAATCTGTGCTGCC 2341  
Dh 2282 ACCTCAGTTTCACTGCTTAAAGCTGCTGCCAGTGCCTTGAATCTGTGCTGCC 2341  
Qy 2342 ATCTATTAACAGATCAAAATTAATAGACCCCGAGTAAAAATTAAGTAGAGCAAGAGGT 2401  
Dh 2342 ATCTATTAACAGATCAAAATTAATAGACCCCGAGTAAAAATTAAGTAGAGCAAGAGGT 2401  
Qy 2402 AGCTTGTTCAAAGATTTTGTGATTTGGGAGCACTGTGATCAAGAGCATCTGT 2461  
Dh 2402 AGCTTGTTCAAAGATTTTGTGATTTGGGAGCACTGTGATCAAGAGCATCTGT 2461  
Qy 2462 TAGTGAGGACACCAAAACCTGTGTGACCGTTTTTTCATGTATGAATTTTGTGTTAGGT 2521  
Dh 2462 TAGTGAGGACACCAAAACCTGTGTGACCGTTTTTTCATGTATGAATTTTGTGTTAGGT 2521

QY 2522 TCGTCTAGTAGCTGTGAGGCTGCTGCTTCTTAGTGGTATGAGAGGAGACATC 2581  
 Db 2522 TCGTCTAGTAGCTGTGAGGCTGCTGCTTCTTAGTGGTATGAGAGGAGACATC 2581  
 QY 2582 TACCAAAATCCATTAGAGTAAACAGCTCTCATGACAGAGGAAACTAATCTCAATGTT 2641  
 Db 2582 TACCAAAATCCATTAGAGTAAACAGCTCTCATGACAGAGGAAACTAATCTCAATGTT 2641  
 QY 2642 TTAAGTATATAAAGCTGTACGCAAGTACTTTGAGCATATAAAAAAAAAAAAAAANG 2701  
 Db 2642 TTAAGTATATAAAGCTGTACGCAAGTACTTTGAGCATATAAAAAAAAAAAAAAANG 2701  
 QY 2702 GCGGCGCGC 2710  
 Db 2702 GCGGCGCGC 2710

RESULT 8  
 US-09-464-231-8  
 ; Sequence 8, Application US/09464231  
 ; Patent No. 6562343  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Levinson, Douglas A.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 ; NUMBER OF SEQUENCES: 37  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036/2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/464,231  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/609,583  
 ; FILING DATE: 01-MAR-1996  
 ; APPLICATION NUMBER: US 08/487,748  
 ; FILING DATE: 07-JUN-1995  
 ; APPLICATION NUMBER: US 08/398,633  
 ; FILING DATE: 03-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Cornuzzi, Laura A.  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 7853-048  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-790-9090  
 ; TELEFAX: 212-869-8864  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2710 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: DNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 40..885  
 ; US-09-464-231-8

Query Match 99.9%; Score 2706.6; DB 4; Length 2710;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGACACCGGCTCCGGATTTTCCCTCCCAAGTACTCATGTTTTCAGGCTTACCTCA 61  
 Db 2 GTGACACCGGCTCCGGATTTTCCCTCCCAAGTACTCATGTTTTCAGGCTTACCTCA 61  
 QY 62 ACTGTGCTCTGCTGTGCTGCACTACTTGTGCAAGGTCATTGGAAGTGTATTAAG 121  
 Db 62 ACTGTGCTCTGCTGTGCTGCACTACTTGTGCAAGGTCATTGGAAGTGTATTAAG 121  
 QY 122 TTGAGGTGGTAAAAATCCCTATCTGCCCCGACGTTAACACTTACTCATTTGGGACAC 181  
 Db 122 TTGAGGTGGTAAAAATCCCTATCTGCCCCGACGTTAACACTTACTCATTTGGGACAC 181  
 QY 182 TTGTCCTATGCTGTGGGCAAGGATTTCTGCTTGTGTCACAGTGTACCAATGATTGC 241  
 Db 182 TTGTCCTATGCTGTGGGCAAGGATTTCTGCTTGTGTCACAGTGTACCAATGATTGC 241  
 QY 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATCCAGACAGATACCAAGCTAAAGGCG 301  
 Db 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATCCAGACAGATACCAAGCTAAAGGCG 301  
 QY 302 ATCTCAACAAAGAGATGTCTCTGATCATTAAGAAATGTGACTGTGATGCCATGGGA 361  
 Db 302 ATCTCAACAAAGAGATGTCTCTGATCATTAAGAAATGTGACTGTGATGCCATGGGA 361  
 QY 362 CCTACTGTGACAGATACAGTTCCTGCTTATGATGATGATGATGATGATGATGATGAT 421  
 Db 362 CCTACTGTGACAGATACAGTTCCTGCTTATGATGATGATGATGATGATGATGATGAT 421  
 QY 422 TAGACATCAAGACAGCCAGGTCCTGCTGACAGTCCAGCTGAGGAGCTTACTACAG 481  
 Db 422 TAGACATCAAGACAGCCAGGTCCTGCTGACAGTCCAGCTGAGGAGCTTACTACAG 481  
 QY 482 CTCTCCCAAGAACCTTAACCAAGAGAAATGTGTCAGACACACACCTGTGACCC 541  
 Db 482 CTCTCCCAAGAACCTTAACCAAGAGAAATGTGTCAGACACACACCTGTGACCC 541  
 QY 542 TCCATATATCAATGACAAACAAATTTTCCATGAGGCTGATGATTAAGACTGTGAG 601  
 Db 542 TCCATATATCAATGACAAACAAATTTTCCATGAGGCTGATGATTAAGACTGTGAG 601  
 QY 602 AAACGATCAGAACTGCTATCCACATTTGAGAGTGTCTGCTGCTGCTGCTGCTGCTG 661  
 Db 602 AAACGATCAGAACTGCTATCCACATTTGAGAGTGTCTGCTGCTGCTGCTGCTGCTG 661  
 QY 662 TTATATGCTGTCTTATATCTTAAATGCTATCTGCTGCTGCTGCTGCTGCTGCTG 721  
 Db 662 TTATATGCTGTCTTATATCTTAAATGCTATCTGCTGCTGCTGCTGCTGCTGCTG 721  
 QY 722 TGAGCCTTATTAACATGAGCACTGCTCCAGAGAGGTTGGCAAAATGACAGAGCAGTCA 781  
 Db 722 TGAGCCTTATTAACATGAGCACTGCTCCAGAGAGGTTGGCAAAATGACAGAGCAGTCA 781  
 QY 782 GGATTCGCTGTGAGGAAATATCTACACATGAGAGAAAGTATGAGTGAATTT 841  
 Db 782 GGATTCGCTGTGAGGAAATATCTACACATGAGAGAAAGTATGAGTGAATTT 841  
 QY 842 CAAATGATTACTACTGCTACGTCACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 901  
 Db 842 CAAATGATTACTACTGCTACGTCACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 901  
 QY 902 ACTTTTAAAGGCTGCTTCAATTTCTGACTTGTGATTTCCCTTTTGGAAAACTATGTG 961  
 Db 902 ACTTTTAAAGGCTGCTTCAATTTCTGACTTGTGATTTCCCTTTTGGAAAACTATGTG 961  
 QY 962 ATATGTCACCTTGAGCACTCATTTGAGAGGTTTGAACACAGCAGCAGCAGCAGCAGCAG 1021  
 Db 962 ATATGTCACCTTGAGCACTCATTTGAGAGGTTTGAACACAGCAGCAGCAGCAGCAGCAG 1021  
 QY 1022 GTTTTCGAGATATTAATCAACAAGGAGTATGACGTATCACTATCTCATTTGAAT 1081  
 Db 1022 GTTTTCGAGATATTAATCAACAAGGAGTATGACGTATCACTATCTCATTTGAAT 1081  
 QY 1082 GCTCATTTTATCCCTGAGTTTCAGGAGTGGATGCTCCACTCCAGAGACTTCAATCATG 1141



Db 1082 GCTCCATTATATCCAGAGTTTCAAGGATCGAATCTCCACTCCAGAGCTTCAATATG 1141  
Qy 1142 CGTGTGAAGCTACTGTCCTTTCATACATTAAGATGTTAGTGTGATGTCCTTGAGA 1201  
Db 1142 CGTGTGAAGCTACTGTCCTTTCATACATTAAGATGTTAGTGTGATGTCCTTGAGA 1201  
Qy 1202 CATAGAGTTTGTGTATATCCGAAAGCTCTGAACAGGTAGGGGGAATAAAGGCTAA 1261  
Db 1202 CATAGAGTTTGTGTATATCCGAAAGCTCTGAACAGGTAGGGGGAATAAAGGCTAA 1261  
Qy 1262 GATAGGAAGGTGGGCTTTTGTGATGTTGGAATAATCTTAAGAAGTTGTAGTTCCT 1321  
Db 1262 GATAGGAAGGTGGGCTTTTGTGATGTTGGAATAATCTTAAGAAGTTGTAGTTCCT 1321  
Qy 1322 AGAGATTTCTGACCTTGAAAGATTAAGAAAAAGCCAGGTGGCATATGCTTAACGATAT 1381  
Db 1322 AGAGATTTCTGACCTTGAAAGATTAAGAAAAAGCCAGGTGGCATATGCTTAACGATAT 1381  
Qy 1382 AACTTGGGAACCTTAGCAGAGAGGTGATTAAGTTCAAGTTCAGCCAGGCTATGCTGTA 1441  
Db 1382 AACTTGGGAACCTTAGCAGAGAGGTGATTAAGTTCAAGTTCAGCCAGGCTATGCTGTA 1441  
Qy 1442 AGACTGTCTGAMCATCCAAAGACGAAATATAACATAGACAGCAGAGGCTGGAGATGA 1501  
Db 1442 AGACTGTCTGAMCATCCAAAGACGAAATATAACATAGACAGCAGAGGCTGGAGATGA 1501  
Qy 1502 GGGTCGGACAGTAGAGGTGATGTTGTATCAAGCAGAGAACTATATTTAGTCGTAGACC 1561  
Db 1502 GGGTCGGACAGTAGAGGTGATGTTGTATCAAGCAGAGAACTATATTTAGTCGTAGACC 1561  
Qy 1562 CCACTGAAAAAGCTAGGCTGTGTAAGCATGCTTGTAAGCTCAAGGAATGGAAGGTAA 1621  
Db 1562 CCACTGAAAAAGCTAGGCTGTGTAAGCATGCTTGTAAGCTCAAGGAATGGAAGGTAA 1621  
Qy 1622 AGGCAACAAGATCCCGGGGGCTTGCGTGCAGCTTAGCGCTAGGCTGATTCGAA 1681  
Db 1622 AGGCAACAAGATCCCGGGGGCTTGCGTGCAGCTTAGCGCTAGGCTGATTCGAA 1681  
Qy 1682 GTCCACAAGAGTCCCTGCTCTCAMAATAGATGAGCTGATCTGCGGCAATGTCATAGG 1741  
Db 1682 GTCCACAAGAGTCCCTGCTCTCAMAATAGATGAGCTGATCTGCGGCAATGTCATAGG 1741  
Qy 1742 GGTGTCT 1801  
Db 1742 GGTGTCT 1801  
Qy 1802 ACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGAT 1861  
Db 1802 ACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGAT 1861  
Qy 1862 TCTCTATATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1921  
Db 1862 TCTCTATATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1921  
Qy 1922 CATGCGACAGCTGCT 1981  
Db 1922 CATGCGACAGCTGCT 1981  
Qy 1982 GAGTTTCAAGAGCTGCT 2041  
Db 1982 GAGTTTCAAGAGCTGCT 2041  
Qy 2042 CGCAAGAGCTGCT 2101  
Db 2042 CGCAAGAGCTGCT 2101  
Qy 2102 TGCTCAAGGGGAATGATCTGAGAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2161  
Db 2102 TGCTCAAGGGGAATGATCTGAGAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2161  
Qy 2162 TGACAGAAAGAACCGGTGATGCTGGAATACGAAGGAGAGCTCTTGCAGCAATCTATA 2221  
Db 2162 TGACAGAAAGAACCGGTGATGCTGGAATACGAAGGAGAGCTCTTGCAGCAATCTATA 2221

Db 2162 TGACAGAAAGAACCGGTGATGCTGGAATACGAAGGAGAGCTCTTGCAGCAATCTATA 2221  
Qy 2222 TAGTCAGAAAATATTTCTTTGGAGAGACGCTGTCACAAATTTGATTTTCAAGCCGCTG 2281  
Db 2222 TAGTCAGAAAATATTTCTTTGGAGAGACGCTGTCACAAATTTGATTTTCAAGCCGCTG 2281  
Qy 2282 ACCTGAGTTTCAATCTGCTTGAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2341  
Db 2282 ACCTGAGTTTCAATCTGCTTGAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2341  
Qy 2342 ATCTATACAGAAATCAATTAATTAATAGACCCGAGTGAAATATTAAGTAGCAGAAAGGT 2401  
Db 2342 ATCTATACAGAAATCAATTAATTAATAGACCCGAGTGAAATATTAAGTAGCAGAAAGGT 2401  
Qy 2402 AGCTTTGTTCAAGATTTTGTGATGTTGGAGAGCAATCTGTATCATCAGAGCATCTGT 2461  
Db 2402 AGCTTTGTTCAAGATTTTGTGATGTTGGAGAGCAATCTGTATCATCAGAGCATCTGT 2461  
Qy 2462 TAGTGAGACACCAAAACCTGTGTGACCGTTTTTTCATGATGATGATGATGATGATGATGAT 2521  
Db 2462 TAGTGAGACACCAAAACCTGTGTGACCGTTTTTTCATGATGATGATGATGATGATGATGAT 2521  
Qy 2522 TGCTTCTAGCTAGCTGTGAGGCTCTGCTTTCTTGTAGTGTGATGAGAGAGACATC 2581  
Db 2522 TGCTTCTAGCTAGCTGTGAGGCTCTGCTTTCTTGTAGTGTGATGAGAGAGACATC 2581  
Qy 2582 TAACAAAATCCTTAGAGATTAACAGCTTCATGACAGAGGAAACTAATCTCAATGTT 2641  
Db 2582 TAACAAAATCCTTAGAGATTAACAGCTTCATGACAGAGGAAACTAATCTCAATGTT 2641  
Qy 2642 TTAAGTAATTAACAGCTGCTGCAAGTACTTTGACATTAATAAAAAAAAAAAAAAAG 2701  
Db 2642 TTAAGTAATTAACAGCTGCTGCAAGTACTTTGACATTAATAAAAAAAAAAAAAAAG 2701  
Qy 2702 GGGGGCGC 2710  
Db 2702 GGGGGCGC 2710

RESULT 9  
US-08-829-525-22  
; Sequence 22, Application US/08829525  
; Patent No. 6084083  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Penite & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/829,525  
; FILING DATE: 28-MAR-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.















GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 23:34:40 ; Search time 871.58 Seconds  
(without alignments)  
10161.809 Million cell updates/sec

Title: US-10-004-633-8

Perfect score: 2710

Sequence: 1 ngtcgaccacgcgcgcga.....aaaaaaaaagcgccgcgc 2710

Scoring table:

IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2706.6	99.9	2710	US-10-252-131-8	Sequence 8, Appli
2	2706.6	99.9	2710	US-10-004-633-8	Sequence 8, Appli
3	2624.6	96.8	2725	US-10-188-012-10	Sequence 12, Appli
4	844.4	31.2	862	US-10-188-012-12	Sequence 22, Appli
5	843	31.1	843	US-10-252-131-32	Sequence 22, Appli
6	843	31.1	843	US-10-004-633-22	Sequence 22, Appli
7	499.8	18.4	2236	US-10-252-131-23	Sequence 0, Appli
8	499.8	18.4	2236	US-10-004-633-23	Sequence 23, Appli
9	485.8	17.9	1116	US-10-188-012-30	Sequence 30, Appli
10	485.8	17.9	1116	US-10-188-012-32	Sequence 32, Appli
11	456	16.8	903	US-10-252-131-37	Sequence 37, Appli
12	456	16.8	903	US-10-004-633-37	Sequence 37, Appli
13	328.8	12.1	2091	US-09-529-063-36	Sequence 36, Appli
14	328.8	12.1	2091	US-10-414-378-36	Sequence 36, Appli
15	289	10.7	555	US-09-529-063-35	Sequence 35, Appli
16	289	10.7	555	US-10-414-378-35	Sequence 35, Appli

17	263.4	9.7	529	13	US-10-040-739-415	Sequence 415, App
18	231.6	8.5	1203	10	US-09-966-546-21	Sequence 21, Appli
19	231.6	8.5	1203	10	US-09-966-545-21	Sequence 21, Appli
20	231.6	8.5	1203	11	US-09-965-212-21	Sequence 21, Appli
21	231.6	8.5	1203	15	US-10-189-940-21	Sequence 21, Appli
22	182.6	6.7	375	15	US-10-189-940-144	Sequence 144, App
23	84.2	3.1	1079	14	US-10-188-012-24	Sequence 24, Appli
24	84.2	3.1	1080	14	US-10-188-012-18	Sequence 18, Appli
25	84.2	3.1	1080	14	US-10-188-012-20	Sequence 20, Appli
26	84.2	3.1	1095	14	US-10-188-012-26	Sequence 26, Appli
27	84.2	3.1	1095	14	US-10-188-012-22	Sequence 22, Appli
28	84.2	3.1	1099	14	US-10-188-012-28	Sequence 28, Appli
29	78.2	2.9	1032	14	US-10-188-012-14	Sequence 14, Appli
30	78.2	2.9	1032	14	US-10-188-012-16	Sequence 16, Appli
31	77.8	2.9	918	14	US-10-188-012-2	Sequence 2, Appli
32	76.8	2.8	849	14	US-10-188-012-4	Sequence 4, Appli
33	76.8	2.8	958	14	US-10-188-012-6	Sequence 6, Appli
34	76.8	2.8	958	14	US-10-188-012-8	Sequence 8, Appli
35	70.8	2.6	398	10	US-09-960-352-7732	Sequence 7732, App
36	70	2.6	418	10	US-09-960-352-1063	Sequence 1063, App
37	65.6	2.4	571	9	US-09-739-907-20	Sequence 20, Appli
38	64.6	2.4	263744	12	US-10-229-834A-6	Sequence 6, Appli
39	64.4	2.4	1156	14	US-10-188-012-34	Sequence 34, Appli
40	64.4	2.4	1156	14	US-10-188-012-36	Sequence 36, Appli
41	64.4	2.4	1282	11	US-09-813-153-68	Sequence 68, Appli
42	64.4	2.4	1380	14	US-10-037-270-59	Sequence 59, Appli
43	63	2.3	382	11	US-08-232-785-348	Sequence 348, App
44	62	2.3	361	11	US-09-232-785-265	Sequence 265, App
45	61.4	2.3	3318	12	US-10-027-632-113926	Sequence 113926, App

#### ALIGNMENTS

RESULT 1  
US-10-252-131-8  
; Sequence 8, Application US/10252131  
; Publication No. US20030158399A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edwards  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Fast-Seq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/252,131  
FILING DATE: 20-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Cortuzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864

TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2710 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 40..885  
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-10-252-131-8  
  
Query Match 99.9%; Score 2706.6; DB 12; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 GTGACCCAGCGCTCCGGATTTCCCTCCCAAGTACATGTTTTCAGGCTTACCTCA 61  
Db 2 GTGACCCAGCGCTCCGGATTTCCCTCCCAAGTACATGTTTTCAGGCTTACCTCA 61  
  
QY 62 ACTGTCCTGCTGCTGCTGCTCACTACTTTCAGAGTCAATGGAAGATTATAG 121  
Db 62 ACTGTCCTGCTGCTGCTGCTCACTACTTTCAGAGTCAATGGAAGATTATAG 121  
  
QY 122 TTGAGGTTGTAATAATGCTTATCTGCTGCACTTACCTTACCTTACCTGAGAC 181  
Db 122 TTGAGGTTGTAATAATGCTTATCTGCTGCACTTACCTTACCTTACCTGAGAC 181  
  
QY 182 TTGTCCTATGTCGTGGGGCAAGGGATTCGTCTGCTGTCAGAGTGAACAAGATTC 241  
Db 182 TTGTCCTATGTCGTGGGGCAAGGGATTCGTCTGCTGTCAGAGTGAACAAGATTC 241  
  
QY 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATCAGACAGTACAGCTAAAGGCG 301  
Db 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATCAGACAGTACAGCTAAAGGCG 301  
  
QY 302 ATCTCAACAAAGAGATGTGTCTCTGATCAATAAAGATGTGACTCTGATGACCATGGGA 361  
Db 302 ATCTCAACAAAGAGATGTGTCTCTGATCAATAAAGATGTGACTCTGATGACCATGGGA 361  
  
QY 362 CCTACGCTGCAAGGATCAATCCCTGCTGCTTATGATGATGATGATGATGATGATGAT 421  
Db 362 CCTACGCTGCAAGGATCAATCCCTGCTGCTTATGATGATGATGATGATGATGATGAT 421  
  
QY 422 TAGACATCAAAAGAGCCCAAGTCACTCAGCTCAGCTGAGCTGAGGAGCTTACTACAG 481  
Db 422 TAGACATCAAAAGAGCCCAAGTCACTCAGCTCAGCTGAGCTGAGGAGCTTACTACAG 481  
  
QY 482 CTTCTCCAGAACCTTAAACGAGAGAAATGTTTCAAGACACAGACACTGTGTGACCC 541  
Db 482 CTTCTCCAGAACCTTAAACGAGAGAAATGTTTCAAGACACAGACACTGTGTGACCC 541  
  
QY 542 TCCATATATACATGGAACAAAATTCCCATGAGGCTGATGAAATTAAGAACTCTGGAG 601  
Db 542 TCCATATATACATGGAACAAAATTCCCATGAGGCTGATGAAATTAAGAACTCTGGAG 601  
  
QY 602 AAACGATCAGAACTGCTATCCACATGAGTGGAGTCTCTGCTGCTGTTGACCTGAC 661  
Db 602 AAACGATCAGAACTGCTATCCACATGAGTGGAGTCTCTGCTGCTGTTGACCTGAC 661  
  
QY 662 TTATCATGTTGTTTAACTCTTAAATGTAATTCCTGTAAGAAAAAAGATTCGAGTT 721  
Db 662 TTATCATGTTGTTTAACTCTTAAATGTAATTCCTGTAAGAAAAAAGATTCGAGTT 721  
  
QY 722 TGAGCCTTATTAACCTGAGCACTTGGCTCAGAGAGGTTGGCAAAAGCGAGCAATCA 781  
Db 722 TGAGCCTTATTAACCTGAGCACTTGGCTCAGAGAGGTTGGCAAAAGCGAGCAATCA 781  
  
QY 782 GGAATCGCTCTGAGAAATATCTACACATCGAGAGAAAGTATATGAAGTGAAGAT 841  
Db 782 GGAATCGCTCTGAGAAATATCTACACATCGAGAGAAAGTATATGAAGTGAAGAT 841

QY 842 CAAATGATCTAATCTGCTACGTCACAGCCAGCAGCAGCTTACCGGCTTGGACTGCC 901  
Db 842 CAAATGATCTAATCTGCTACGTCACAGCCAGCAGCAGCTTACCGGCTTGGACTGCC 901  
  
QY 902 ACTTTAAAGGCTCGGCTTCACTTTCGATCTTGTATTTCCCTTTTGGAAAACTATG 961  
Db 902 ACTTTAAAGGCTCGGCTTCACTTTCGATCTTGTATTTCCCTTTTGGAAAACTATG 961  
  
QY 962 ATATGTCATTGAGCACTCATTTGAGAGTTTGCACACAGCCACTGAGAAAAAGTTCCA 1021  
Db 962 ATATGTCATTGAGCACTCATTTGAGAGTTTGCACACAGCCACTGAGAAAAAGTTCCA 1021  
  
QY 1022 GTTTTTCGGGATTAATTAATCAACAAGGAGTTCACCTGTAATCTACTACTTGAAT 1081  
Db 1022 GTTTTTCGGGATTAATTAATCAACAAGGAGTTCACCTGTAATCTACTACTTGAAT 1081  
  
QY 1082 GCTCCATTTTATCCCTGAGTTTCAGGGATCGGATCTCCCACTTCAGAGACTTCAATCAT 1141  
Db 1082 GCTCCATTTTATCCCTGAGTTTCAGGGATCGGATCTCCCACTTCAGAGACTTCAATCAT 1141  
  
QY 1142 CGTTTGAAGCTCACTGCTGCTTTCATACATTAAGAAATGTTAGTGTCTTGAAG 1201  
Db 1142 CGTTTGAAGCTCACTGCTGCTTTCATACATTAAGAAATGTTAGTGTCTTGAAG 1201  
  
QY 1202 CATAGAGTTTGTGTATATCCGCAAAAGTCTCTGAAACAGTGAAGGGGCTAA 1261  
Db 1202 CATAGAGTTTGTGTATATCCGCAAAAGTCTCTGAAACAGTGAAGGGGCTAA 1261  
  
QY 1262 GATAGAAAGTGCGGCTTGTGATGTTGAAAAATCTTAAAGAAATGTTGATCTTTTC 1321  
Db 1262 GATAGAAAGTGCGGCTTGTGATGTTGAAAAATCTTAAAGAAATGTTGATCTTTTC 1321  
  
QY 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGATATGCTTAAACAGATAT 1381  
Db 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGATATGCTTAAACAGATAT 1381  
  
QY 1382 AACTTGGGAACCTTAAAGGAGAGGATGATTAAGTCAAGTCAAGGCTATGCTGAT 1441  
Db 1382 AACTTGGGAACCTTAAAGGAGAGGATGATTAAGTCAAGTCAAGGCTATGCTGAT 1441  
  
QY 1442 AGACTGCTCAGCATCAAGAGCAAGAAATTAACATAGAGACAGAGAGGCTGAGATGA 1501  
Db 1442 AGACTGCTCAGCATCAAGAGCAAGAAATTAACATAGAGACAGAGAGGCTGAGATGA 1501  
  
QY 1502 GGTTCGAGCAAGTGAAGTGTATGTAACAAGACAGAGAACTATATTTGATCTGAGACC 1561  
Db 1502 GGTTCGAGCAAGTGAAGTGTATGTAACAAGACAGAGAACTATATTTGATCTGAGACC 1561  
  
QY 1562 CCAATGAAAAAGCTAGGCTGCTGTAAGAGATGCTTGAAGCTCAAGAGATGAGAGTTAA 1621  
Db 1562 CCAATGAAAAAGCTAGGCTGCTGTAAGAGATGCTTGAAGCTCAAGAGATGAGAGTTAA 1621  
  
QY 1622 AGGCAACAAGATCCCGGGGCTTGGCTGACAGTCAAGTCAAGCTTAAAGTGTGATTC 1681  
Db 1622 AGGCAACAAGATCCCGGGGCTTGGCTGACAGTCAAGTCAAGCTTAAAGTGTGATTC 1681  
  
QY 1682 GTCCACAAGAGTCCCTGTCTCAMAATGTAAGTGTGATGATGATGATGATGATGAT 1741  
Db 1682 GTCCACAAGAGTCCCTGTCTCAMAATGTAAGTGTGATGATGATGATGATGATGAT 1741  
  
QY 1742 GGTTCCTCTCTCTCACAAGAGATGACATGACATGACATGACATGACATGACATGAC 1801  
Db 1742 GGTTCCTCTCTCTCACAAGAGATGACATGACATGACATGACATGACATGACATGAC 1801  
  
QY 1802 ACACACACACACACACACACACACACATGAAATGAAGTTCTCTGCTGCTGCTAC 1861  
Db 1802 ACACACACACACACACACACACACACATGAAATGAAGTTCTCTGCTGCTGCTAC 1861  
  
QY 1862 TCTCTAATCATGATCTTACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
Db 1862 TCTCTAATCATGATCTTACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921



```

QY 1922 CATGGCAGAGCAGTCAGTAATTTATTCAGACCTCAGAGGCTGGAGCAGAGCGTGA 1981
DB 1922 CATGGCAGAGCAGTCAGTAATTTATTCAGACCTCAGAGGCTGGAGCAGAGCGTGA 1981
QY 1982 GAGTTCAGAGCAGCTGTGCCCACTGCGACACTCTTTTACACAGAAAAAGTTACC 2041
DB 1982 GAGTTCAGAGCAGCTGTGCCCACTGCGACACTCTTTTACACAGAAAAAGTTACC 2041
QY 2042 CGCAGACAGCTGTGTCTGTAAAGGAAACCTGGGAAAGCAACTTGACTGTGTG 2101
DB 2042 CGCAGACAGCTGTGTCTGTAAAGGAAACCTGGGAAAGCAACTTGACTGTGTG 2101
QY 2102 TGCTCAGAGGGAATCTGACTCAGACCACTTCTCCATTCCTGAGGAACTGAGCTGTTC 2161
DB 2102 TGCTCAGAGGGAATCTGACTCAGACCACTTCTCCATTCCTGAGGAACTGAGCTGTTC 2161
QY 2162 TGACGAGAGAACACCGGTGACTGGGACATACGAGGAGAGCTCTTGACAGCAATCTATA 2221
DB 2162 TGACGAGAGAACACCGGTGACTGGGACATACGAGGAGAGCTCTTGACAGCAATCTATA 2221
QY 2222 TAGTACGAGAAATATTTCTTGGGAGGACAGTGTCCCAATTTGATTCAGACCGGTG 2281
DB 2222 TAGTACGAGAAATATTTCTTGGGAGGACAGTGTCCCAATTTGATTCAGACCGGTG 2281
QY 2282 ACCTCAGTTTCACTGTGCTTACAGCTGCTGCCAGTGTCTGATCTGTGCTGTCCC 2341
DB 2282 ACCTCAGTTTCACTGTGCTTACAGCTGCTGCCAGTGTCTGATCTGTGCTGTCCC 2341
QY 2342 ATCTATACAGAAATCAATTAATTAATAGACCCGAGTGAAATATTAAGTACAGAAAGGT 2401
DB 2342 ATCTATACAGAAATCAATTAATTAATAGACCCGAGTGAAATATTAAGTACAGAAAGGT 2401
QY 2402 ACCTTGTTCAGAAATTTTTTGGCAATGGGAGGCACTGTGTACATACAGACATCTGT 2461
DB 2402 ACCTTGTTCAGAAATTTTTTGGCAATGGGAGGCACTGTGTACATACAGACATCTGT 2461
QY 2462 TAGTACGAGACACCAAAACCTGTGTACCTGTTTTCATGATGAAATTTGTTGTTAGGT 2521
DB 2462 TAGTACGAGACACCAAAACCTGTGTACCTGTTTTCATGATGAAATTTGTTGTTAGGT 2521
QY 2522 TGCTTCTAGCTAGCTGTGAGGTCTGTGCTTTCTTGTAGTGGGTATGGAAGGAGCAATC 2581
DB 2522 TGCTTCTAGCTAGCTGTGAGGTCTGTGCTTTCTTGTAGTGGGTATGGAAGGAGCAATC 2581
QY 2582 TAAACAAATTCATTAAGATTAACAGCTCAGACGAGAGGAAACTATCTCAATGTT 2641
DB 2582 TAAACAAATTCATTAAGATTAACAGCTCAGACGAGAGGAAACTATCTCAATGTT 2641
QY 2642 TTAAGTATATAAACTGTACTGCAAAAGTACTTGTGACATATAAAAAAAAAAAG 2701
DB 2642 TTAAGTATATAAACTGTACTGCAAAAGTACTTGTGACATATAAAAAAAAAAAG 2701
QY 2702 GGGGGGGG 2710
DB 2702 GGGGGGGG 2710

```

```

RESULT 2
US-10-004-633-8
; Sequence 8, Application US/10004633
; Publication No. US20030069196A1
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; APPLICANT: Lloyd, Clare M.
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-125
; CURRENT APPLICATION NUMBER: US/10/004,633
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583

```

```

; PRIOR FILING DATE: EARLIER FILING DATE: 1996-03-01
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/487,748
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/398,633
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 2710
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all "n" positions
; OTHER INFORMATION: n=a, c, g, or t
US-10-004-633-8

Query Match          99.9%; Score 2706.6; DB 14; Length 2710;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 2 GTGACCCACGCGTCCGATTTCCCTCCCAAGTACTAGTTTCAGGCTTTACCTCA 61
DB 2 GTGACCCACGCGTCCGATTTCCCTCCCAAGTACTAGTTTCAGGCTTTACCTCA 61
QY 62 ACTGTCCTGCTGCTGCTGCAACTACTTGCAGAGGCACTTGAAGATGTTAAG 121
DB 62 ACTGTCCTGCTGCTGCTGCAACTACTTGCAGAGGCACTTGAAGATGTTAAG 121
QY 122 TTGAGGTGGTAAATGCTTATCTGCTGAGTTACACTTACTACATCTGGGACAC 181
DB 122 TTGAGGTGGTAAATGCTTATCTGCTGAGTTACACTTACTACATCTGGGACAC 181
QY 182 TTGTCCTATGCTGTGGGCAAGGATCTGTCTTGTGCTGACAGTACCAATGAGTTC 241
DB 182 TTGTCCTATGCTGTGGGCAAGGATCTGTCTTGTGCTGACAGTACCAATGAGTTC 241
QY 242 TCAGACATGATGAAGAAATGATATATCAAGATCAAGAGATCAAGCTTAAAGGGG 301
DB 242 TCAGACATGATGAAGAAATGATATATCAAGATCAAGAGATCAAGCTTAAAGGGG 301
QY 302 ATCTCAACAAAGAGATGCTCTGATCATTAAGATGTGACTGTGATGACATGGGA 361
DB 302 ATCTCAACAAAGAGATGCTCTGATCATTAAGATGTGACTGTGATGACATGGGA 361
QY 362 CCTACTGCTGAGATACAGTTCCCTGTCTTATGATGATTAATAATTAGACTGAAT 421
DB 362 CCTACTGCTGAGATACAGTTCCCTGTCTTATGATGATTAATAATTAGACTGAAT 421
QY 422 TAGACATCAAGACACCAAGGCTCACTCAGCTCAAGCTGACATGGGACTTACTACG 481
DB 422 TAGACATCAAGACACCAAGGCTCACTCAGCTCAAGCTGACATGGGACTTACTACG 481
QY 482 CTTCTCAGAAACCTTAACACGAGAGAAATGTTTACAGACAGACACTGTGACCC 541
DB 482 CTTCTCAGAAACCTTAACACGAGAGAAATGTTTACAGACAGACACTGTGACCC 541
QY 542 TCCATTAATTAACATGAACAAATTTCCACATGGGCTGATGAATTAAGACTCTGAG 601
DB 542 TCCATTAATTAACATGAACAAATTTCCACATGGGCTGATGAATTAAGACTCTGAG 601
QY 602 AAACGATCAAGACCTGCTATCCACATTTGAGAGGAGTCTGTGCTGAGTCCCTGGC 661
DB 602 AAACGATCAAGACCTGCTATCCACATTTGAGAGGAGTCTGTGCTGAGTCCCTGGC 661
QY 662 TTATCATGTTGTCTTAATCTTAATGATATCTCTGTAAGAAAAAGAGTATGAGTT 721
DB 662 TTATCATGTTGTCTTAATCTTAATGATATCTCTGTAAGAAAAAGAGTATGAGTT 721
QY 722 TGAGCTTATTAACACTGACCACTTGTCTTCAAGAGGCTTGCAATGACAGGACGTCA 781
DB 722 TGAGCTTATTAACACTGACCACTTGTCTTCAAGAGGCTTGCAATGACAGGACGTCA 781

```

QY	782	GGATTGCGCTGAGGAGAAATATCTACACATCGAGGAAAGATATATAATGAGAAATT	841
Db	782	GGATTGCGCTGAGGAGAAATATCTACACATCGAGGAGAAAGATATATAATGAGAAATT	841
QY	842	CAATAGAGTACTACTGCTACGTCGACAGCCAGCAGCCATCCTGACCGCTCTGACCTGCC	901
Db	842	CAATAGAGTACTACTGCTACGTCGACAGCCAGCAGCCATCCTGACCGCTCTGACCTGCC	901
QY	902	ACTTTTAAAGGCTGCGCTTCATTTCTGACTTTGGATTTCCCTTTKGGAAAATGATGTG	961
Db	902	ACTTTTAAAGGCTGCGCTTCATTTCTGACTTTGGATTTCCCTTTKGGAAAATGATGTG	961
QY	962	ATATGTCACTTGGCAACCTCATTTGAGGTTCTGACCAACAGCCACTGAGAAAAGATTCCA	1021
Db	962	ATATGTCACTTGGCAACCTCATTTGAGGTTCTGACCAACAGCCACTGAGAAAAGATTCCA	1021
QY	1022	GTTTTCTGCGGGAATTAATTAATCTACAGAGGGGATTCGACTGTAACTCATCTACATTTGAAT	1081
Db	1022	GTTTTCTGCGGGAATTAATTAATCTACAGAGGGGATTCGACTGTAACTCATCTACATTTGAAT	1081
QY	1082	GCTCCATTTTATCCCTGAGTTTCAGGATTCGGATCTCCACTCCAGAGACTTCATCATG	1141
Db	1082	GCTCCATTTTATCCCTGAGTTTCAGGATTCGGATCTCCACTCCAGAGACTTCATCATG	1141
QY	1142	CGTGTGAAAGCTCACTCGCTCTTATATCATTAGAAATGTTAGTGTATGCTTTTGAGA	1201
Db	1142	CGTGTGAAAGCTCACTCGCTCTTATATCATTAGAAATGTTAGTGTATGCTTTTGAGA	1201
QY	1202	CATAGAGGTTTGTGTATATCCGCAAGGCTCCTGAAACAGGTAGGGGGAAATTAAGGGCTPA	1261
Db	1202	CATAGAGGTTTGTGTATATCCGCAAGGCTCCTGAAACAGGTAGGGGGAAATTAAGGGCTPA	1261
QY	1262	GATAGGAAGGTGCGGTCTTTTGTGTGTAAGTGTGAAAACTTAAAGAGTTGGTAGCTTTTCT	1321
Db	1262	GATAGGAAGGTGCGGTCTTTTGTGTGTAAGTGTGAAAACTTAAAGAGTTGGTAGCTTTTCT	1321
QY	1322	AGAGATTTCTGACTTTGAAAGTTTAAGAAAAAGCAGGTGCGATATGCTTAAACGAGAT	1381
Db	1322	AGAGATTTCTGACTTTGAAAGTTTAAGAAAAAGCAGGTGCGATATGCTTAAACGAGAT	1381
QY	1382	AACTTGGGAACTTTAGGACGAGAGGTGTAAAGTTCAAGTCAAGCCAGGGCTATGCTGTATA	1441
Db	1382	AACTTGGGAACTTTAGGACGAGAGGTGTAAAGTTCAAGTCAAGCCAGGGCTATGCTGTATA	1441
QY	1442	AGACTGTCTCAGCATTCGAAAGACGAAAAATTAACATAGAGACAGCAGGAGGCTGAGATGA	1501
Db	1442	AGACTGTCTCAGCATTCGAAAGACGAAAAATTAACATAGAGACAGCAGGAGGCTGAGATGA	1501
QY	1502	GGCTGCGACAGTGAAGGTGATGTGTAAAGGACAGAGAAATCTATAATTGATCGTAGACC	1561
Db	1502	GGCTGCGACAGTGAAGGTGATGTGTAAAGGACAGAGAAATCTATAATTGATCGTAGACC	1561
QY	1562	CCACATGAAAAAGCTAGGCTTGTAAGACATGCTTGTAGACTCAAGAGANTGAGAGGTAA	1621
Db	1562	CCACATGAAAAAGCTAGGCTTGTAAGACATGCTTGTAGACTCAAGAGANTGAGAGGTAA	1621
QY	1622	AGGCACACAGATCCCGGGGCTTGCGTGACGTCAAGCTTAAGCTTAAGGTGAGTGTCCAA	1681
Db	1622	AGGCACACAGATCCCGGGGCTTGCGTGACGTCAAGCTTAAGCTTAAGGTGAGTGTCCAA	1681
QY	1682	GTCACCAAGAGTCCCTGTCTCAGATGATGATGATGATCTGAGCGCATGTCCATGGG	1741
Db	1682	GTCACCAAGAGTCCCTGTCTCAGATGATGATGATGATCTGAGCGCATGTCCATGGG	1741
QY	1742	GGTGTGCTCTCTCTCTCAGAAAGACATGACATATWCCCTGACACACACACACACACAC	1801
Db	1742	GGTGTGCTCTCTCTCTCAGAAAGACATGACATATWCCCTGACACACACACACACACAC	1801
QY	1802	ACACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGATG	1861
Db	1802	ACACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGATG	1861
QY	1862	TCTCTATTAACATGATCTCTACAGGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAC	1921

Db	1862	TCTCTAATAATGATATCTTACAGGACTCTCTCTGCTCTGTAAAGACATGATGGAG	1921
Oy	1922	CATGGCAGACAGTCCAGTAATTTATTCACGACTCAGAAAGGCTGGAGCGAAAGCCGTGA	1981
Db	1922	CATGGCAGACAGTCCAGTAATTTATTCACGACTCAGAAAGGCTGGAGCGAAAGCCGTGA	1981
Oy	1982	GAGTTTCAGAGACATGTGCCCCAACACTGCGCAGACTCTTTCTTACACAAGAAAAGGTATCC	2041
Db	1982	GAGTTTCAGAGACATGTGCCCCAACACTGCGCAGACTCTTTCTTACACAAGAAAAGGTATCC	2041
Oy	2042	CGCAAGCAGCCCTGCTGTCTGTAAAAAGAAAACCTGCGAAAGCGAACTTTGACTGTTGTG	2101
Db	2042	CGCAAGCAGCCCTGCTGTCTGTAAAAAGAAAACCTGCGAAAGCGAACTTTGACTGTTGTG	2101
Oy	2102	TGCTCAAGGGGAACGTGACTCAGACCAATTCCTCAGACTCTGAGAGAAACTGAGACTGTTC	2161
Db	2102	TGCTCAAGGGGAACGTGACTCAGACCAATTCCTCAGACTCTGAGAGAAACTGAGACTGTTC	2161
Oy	2162	TGACAGAAAGAACACCGGTGACTGGGACATACGAAAGGCAAGCCTTGCAGCAATCTATA	2221
Db	2162	TGACAGAAAGAACACCGGTGACTGGGACATACGAAAGGCAAGCCTTGCAGCAATCTATA	2221
Oy	2222	TAGTCACCAAAATATTTCTTTGGAGAGCACTGCTGCACCAATTTGATTTCCAGCCGGTGG	2281
Db	2222	TAGTCACCAAAATATTTCTTTGGAGAGCACTGCTGCACCAATTTGATTTCCAGCCGGTGG	2281
Oy	2282	ACCTCACTTTCATCTGGCTTACAGCTGCTGCCCCAGTCCCTTGATCTGATGCTGCTCCC	2341
Db	2282	ACCTCACTTTCATCTGGCTTACAGCTGCTGCCCCAGTCCCTTGATCTGATGCTGCTCCC	2341
Oy	2342	ATCTATTAACGAATCAAATTAATAGACCCCGAGTAAAAATTTAAGTGAACGAAAGGT	2401
Db	2342	ATCTATTAACGAATCAAATTAATAGACCCCGAGTAAAAATTTAAGTGAACGAAAGGT	2401
Oy	2402	AGCTTTGTCAAAAGATTTTTTTTGACTGGGAGCACTGATCATACAGAGCATCTGT	2461
Db	2402	AGCTTTGTCAAAAGATTTTTTTTGACTGGGAGCACTGATCATACAGAGCATCTGT	2461
Oy	2462	TAGTGAAGAACCCAAACCTGTGGTACCGTTTTTTCATGATGAATTTGTGTTTAGT	2521
Db	2462	TAGTGAAGAACCCAAACCTGTGGTACCGTTTTTTCATGATGAATTTGTGTTTAGT	2521
Oy	2522	TGCTTCAGTACGTAGCTGAGAGCTCTGCGCTTCTTACGTGGGTATGGAAGGAGACATC	2581
Db	2522	TGCTTCAGTACGTAGCTGAGAGCTCTGCGCTTCTTACGTGGGTATGGAAGGAGACATC	2581
Oy	2582	TAACAAAATCATTAGAGATPAACAGCTCTCATGCAAGGAGAAAACCTAATCTCAATGTT	2641
Db	2582	TAACAAAATCATTAGAGATPAACAGCTCTCATGCAAGGAGAAAACCTAATCTCAATGTT	2641
Oy	2642	TTAAAGTAATAAATCTGTACTGCGAAAGTACTTTAGCATTAATAAAAAAAAAAAAAAAG	2701
Db	2642	TTAAAGTAATAAATCTGTACTGCGAAAGTACTTTAGCATTAATAAAAAAAAAAAAAAAG	2701
Oy	2702	GGCGGCGCG 2710	
Db	2702	GGCGGCGCG 2710	
RESULT 3			
US-10-188-012-10			
; Sequence 10, Application US/10188012			
; Publication No. US20030124114A1			
; GENERAL INFORMATION:			
; APPLICANT: McIntire, Jennifer Jones			
; APPLICANT: Umetsu, Dale T.			
; APPLICANT: Dekruyff, Rosemarie			
; APPLICANT: Kuchroo, Vijay			
; APPLICANT: Freeman, Gordon J.			
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of			
; FILE REFERENCE: STAN-235			

; CURRENT APPLICATION NUMBER: US/10/188,012  
; CURRENT FILING DATE: 2002-07-01  
; PRIOR APPLICATION NUMBER: 60/302,344  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 2725  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-188-012-10

Query Match 96.8%; Score 2624.6; DB 14; Length 2725;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 2655; Conservative 6; Mismatches 20; Indels 2; Gaps 2;

QY 18 GGATTTCCCTCCCAAGTACTCATGTTTCAGGCTTACCTCAACTGTGCTGCTGCT  
DB 45 GGATTTCCCTCCCAAGTACTCATGTTTCAGGCTTACCTCAACTGTGCTGCTGCT  
QY 78 GCTGCACTACTACTTGAAGGTCATTTGGAGAATGGTTATTAAGTTGAGTTGTTGA  
DB 105 GCTGCACTACTACTTGAAGGTCATTTGGAGAATGGTTATTAAGTTGAGTTGTTGA  
QY 138 TGCCATATCGCCCTGCAAGTTCACCTTACCTTACCTGAGACACTTGTGCTATGCTG  
DB 165 TGCCATATCGCCCTGCAAGTTCACCTTACCTTACCTGAGACACTTGTGCTATGCTG  
QY 198 GGGCAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  
DB 225 GGGCAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  
QY 258 AAATGTGACATTAAGAAATCCAGAGATACAGAGATACAGAGATACAGAGATACAGAG  
DB 265 AAATGTGACATTAAGAAATCCAGAGATACAGAGATACAGAGATACAGAGATACAGAG  
QY 318 TGTGCTCTGATCATTAAGAAATGATGATGATGATGATGATGATGATGATGATGAT  
DB 345 TGTGCTCTGATCATTAAGAAATGATGATGATGATGATGATGATGATGATGATGAT  
QY 378 ACAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  
DB 405 ACAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  
QY 438 CAAGGTCACTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG  
DB 465 CAAGGTCACTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG  
QY 498 AACCAAGGAGAAATGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT  
DB 525 AACCAAGGAGAAATGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT  
QY 558 AACCAAAATTTTCCATGAGGCTGATGAAATTAAGAACTTCTGGAAGAACGATCAGAACTGC  
DB 585 AACCAAAATTTTCCATGAGGCTGATGAAATTAAGAACTTCTGGAAGAACGATCAGAACTGC  
QY 618 TATCAACATTTGAGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  
DB 645 TATCAACATTTGAGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  
QY 678 AATCTTAATGATTTCTGTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG  
DB 705 AATCTTAATGATTTCTGTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG  
QY 738 GGCCTAATTTGCTGCAAGGAGGTTGCAAAATCAGAGAGAGTCAAGATTCGCTGAGAGA  
DB 765 GGCCTAATTTGCTGCAAGGAGGTTGCAAAATCAGAGAGAGTCAAGATTCGCTGAGAGA  
QY 798 AAATATCTACACATCAGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT  
DB 825 AAATATCTACACATCAGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT  
QY 858 CTACGTCAACAGCCAGCCATCTGTAACCGCTCTGACTGCACTTTTAAGGCTGCG 917

DB 885 CTACGTCAACAGCCAGCCATCTGTAACCGCTCTGACTGCACTTTTAAGGCTGCG 944  
QY 918 CTTCATTTTCTGACTTTGGTATTTCCCTTTTGTGGAATATGATGATGATGATGATGATGAT  
DB 945 CTTCATTTTCTGACTTTGGTATTTCCCTTTTGTGGAATATGATGATGATGATGATGATGAT  
QY 978 CCTCATTTGAGGTTCTGCAACAGCCATCTGTAACCGCTCTGACTGCACTTTTAAGGCTGCG  
DB 1004 CCTCATTTGAGGTTCTGCAACAGCCATCTGTAACCGCTCTGACTGCACTTTTAAGGCTGCG  
QY 1038 TTAATCAACAGGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT  
DB 1064 TTAATCAACAGGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT  
QY 1098 GAGTTTCAGGAGTGGATTCCTCCACTCCAGAGACTTCAATCATGCTGCTGGAAGTCACT  
DB 1124 GAGTTTCAGGAGTGGATTCCTCCACTCCAGAGACTTCAATCATGCTGCTGGAAGTCACT  
QY 1158 GGTGCTTTCATACATTAAGAAATGGTTAGTGTGATGATGATGATGATGATGATGATGAT  
DB 1184 GGTGCTTTCATACATTAAGAAATGGTTAGTGTGATGATGATGATGATGATGATGATGAT  
QY 1218 ATATCCGCAAGGCTCTGTAACAGGATGAGGAGATTAAGGCTTAAGATGAGAGTGCCT  
DB 1244 ATATCCGCAAGGCTCTGTAACAGGATGAGGAGATTAAGGCTTAAGATGAGAGTGCCT  
QY 1278 CTTTGTGATGTTGGAATCTTTAAAGATTTGATGATGATGATGATGATGATGATGATGAT  
DB 1304 CTTTGTGATGTTGGAATCTTTAAAGATTTGATGATGATGATGATGATGATGATGATGAT  
QY 1338 GAAAGATTAAGAAATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT  
DB 1363 GAAAGATTAAGAAATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT  
QY 1398 GCAGGAGGATTAAGTTCAAGGTCAGGAGGATGATGATGATGATGATGATGATGATGAT  
DB 1423 GCAGGAGGATTAAGTTCAAGGTCAGGAGGATGATGATGATGATGATGATGATGATGAT  
QY 1458 CAAGAAGAAATTAAGATTAAGAGACAGAGAGGCTGGAATGAGGCTGGAAGTGAAG  
DB 1483 CAAGAAGAAATTAAGATTAAGAGACAGAGAGGCTGGAATGAGGCTGGAAGTGAAG  
QY 1518 TGCATTTGTAACAGCAAGGAGATCTATTTTGAATGATGATGATGATGATGATGATGAT  
DB 1543 TGCATTTGTAACAGCAAGGAGATCTATTTTGAATGATGATGATGATGATGATGATGAT  
QY 1578 GGCCTGTAAGCAATGCTTTGTAAGTCAAGAGATGAGAGGTTAAAGGACACAGATCCC  
DB 1603 GGCCTGTAAGCAATGCTTTGTAAGTCAAGAGATGAGAGGTTAAAGGACACAGATCCC  
QY 1638 CGGGGCTTGCCTGCAAGTCAAGCTTAAGGTTGCTGATGATGATGATGATGATGATGAT  
DB 1663 CGGGGCTTGCCTGCAAGTCAAGCTTAAGGTTGCTGATGATGATGATGATGATGATGAT  
QY 1698 GTCCTAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG  
DB 1723 GTCCTAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG  
QY 1758 CAGAAGAGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATGAC  
DB 1783 CAGAAGAGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATGAC  
QY 1818 ACACACACACACATGAAATGAAAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  
DB 1843 ACACACACACACATGAAATGAAAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  
QY 1878 CTCTACAGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  
DB 1903 CTCTACAGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  
QY 1938 AGTAATTAATTCAGACACTCAGAAAGGCTGAGAGAGACGCTGAGAGATTCAGAGACACTG 1997

Db 1963 AGTAATTAATTCAGACATCTGAGAAAGCTGAGAGAAAGCTGAGAGATTTCAGAGACATG 2022  
QY 1998 TGGCCAACTGCGCAAGCTCTTCTTACACAGAAAAAGTTACCCGCAAGACAGCTGCTG 2057  
Db 2023 TGGCCAACTGCGCAAGCTCTTCTTACACAGAAAAAGTTACCCGCAAGACAGCTGCTG 2082  
QY 2058 TCTGTAAAGGAAACCTGCGAAAGGCAAACTTGTGCTGTGTGCTGCAAGGGAACG 2117  
Db 2083 TCTGTAAAGGAAACCTGCGAAAGGCAAACTTGTGCTGTGTGCTGCAAGGGAACG 2142  
QY 2118 ACTCAGACAACCTCTCCATCTCTGAGAGAACTGAGAGCTGTTCTGACAGAAACACAC 2177  
Db 2143 ACTCAGACAACCTCTCCATCTCTGAGAGAACTGAGAGCTGTTCTGACAGAAACACAC 2202  
QY 2178 GGTGACTGGGACATACGAGAGGAGAGAGCTTTCAGAGCAATCTATATAGTCAGAAAAATTT 2237  
Db 2203 GGTGACTGGGACATACGAGAGGAGAGAGCTTTCAGAGCAATCTATATAGTCAGAAAAATTT 2262  
QY 2238 CTTTGGAGAGACAGTGTGTCACCAAAATTGATTTCCAGCCGGTGGACCTGATTCATCTG 2297  
Db 2263 CTTTGGAGAGACAGTGTGTCACCAAAATTGATTTCCAGCCGGTGGACCTGATTCATCTG 2322  
QY 2298 GCTTACAGCTGCTGCGCCAGAGCTTGTATCTGTGCTGCTCCCATCTATACAGATCA 2357  
Db 2323 GCTTACAGCTGCTGCGCCAGAGCTTGTATCTGTGCTGCTCCCATCTATACAGATCA 2382  
QY 2358 AATTAAATAGACCCCGAGTGAATAATTAAAGTACAGAAAGTAGTACCTTTGTTCAAGAT 2417  
Db 2383 AATTAAATAGACCCCGAGTGAATAATTAAAGTACAGAAAGTAGTACCTTTGTTCAAGAT 2442  
QY 2418 TTTTTCGATTTGGGAGCACTGTGATCATCAGAGGACATCTGTGTGTGAGACACCAAA 2477  
Db 2443 TTTTTCGATTTGGGAGCACTGTGATCATCAGAGGACATCTGTGTGTGAGACACCAAA 2502  
QY 2478 ACCGTGTGACCGTTTTTCATGATGATTAATTTTGTGTAGGTGCTTCTAGCTAGCTG 2537  
Db 2503 ACCGTGTGACCGTTTTTCATGATGATTAATTTTGTGTAGGTGCTTCTAGCTAGCTG 2562  
QY 2538 TGGAGTCTGCTGCTTTTCTTATGTTGGTGTATGGAAGGAGACCATCTTAAACAAATCCATTAG 2597  
Db 2563 TGGAGTCTGCTGCTTTTCTTATGTTGGTGTATGGAAGGAGACCATCTTAAACAAATCCATTAG 2622  
QY 2598 AGATTAACAGCTCTCAGTCAAGAGGAAAACTAATCTCAAAATGTTTAAATTAATAACT 2657  
Db 2623 AGATTAACAGCTCTCAGTCAAGAGGAAAACTAATCTCAAAATGTTTAAATTAATAACT 2682  
QY 2658 GTACTGGCAAAAGTACTTGTGAGCATTAATAAAAAAAAAAAAAAAAAAAAAA 2700  
Db 2683 GTACTGGCAAAAGTACTTGTGAGCATATTTAATAAAAAAAAAAAAAAAAAAAAAA 2725

RESULT 4  
US-10-188-012-12  
; Sequence 12, Application US/10188012  
; Publication No. US20030124114A1  
; GENERAL INFORMATION:  
; APPLICANT: McInchre, Jennifer Jones  
; APPLICANT: Umetsu, Dale T.  
; APPLICANT: Dekruyf, Rosemarie  
; APPLICANT: Kuchoo, Vijay  
; APPLICANT: Freeman, Gordon J.  
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
; FILE REFERENCE: STAN-235  
; CURRENT APPLICATION NUMBER: US/10/188,012  
; PRIORITY FILING DATE: 2002-07-01  
; PRIOR APPLICATION NUMBER: 60/302,344  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 862  
; TYPE: DNA

; ORGANISM: Mus musculus  
US-10-188-012-12  
Query Match 31.2%; Score 844.4; DB 14; Length 862;  
Best Local Similarity 98.7%; Pred. No. 2,1e-245;  
Matches 851; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 24 CCCCTCCCAAGTACTCATGTTTTCAGGTTTACCTCTCAACCTGTGTCTGTCTGTGCA 83  
Db 1 CCCCTCCCAAGTACTCATGTTTTCAGGTTTACCTCTCAACCTGTGTCTGTCTGTGCA 60  
QY 84 ACTACTACTTGGAGGTCATTTGGAAGATGTTATAGGTTGAGTTGTTAAATATCTTA 143  
Db -61 ACTACTACTTGGAGGTCATTTGGAAGATGTTATAGGTTGAGTTGTTAAATATCTTA 120  
QY 144 TCTGCCCTGCAAGTACATCTTACCTACATCTTGGACACTTGTGCTATGCTGGGCA 203  
Db 121 TCTGCCCTGCAAGTACATCTTACCTACATCTTGGGCACTTGTGCTATGCTGGGCA 180  
QY 204 GGGATTCCTGCTGTGTGTCAGTGTACCAATGATGCTCAGAACTGATGAAAGAAATGT 263  
Db 181 GGGATTCCTGCTGTGTGTCAGTGTACCAATGATGCTCAGAACTGATGAAAGAAATGT 240  
QY 264 GACATATCAGAAATCCAGAGATACAGCTTAAAGGCGATCTCAACAAAGAGATGTGTC 323  
Db 241 GACATATCAGAAATCCAGAGATACAGCTTAAAGGCGATCTCAACAAAGAGAGATGTGTC 300  
QY 324 TCTGATCATTAAGATGTGACTCTGATGATCAGATGGACCTATCTGTGAGATACAGTT 383  
Db 301 TCTGATCATTAAGATGTGACTCTGATGATCAGATGGACCTATCTGTGAGATACAGTT 360  
QY 384 CCTGTCTTATGATGATTAATAATTAAGTAACTGAAATTAAGTAAAGTAAAGTAAAGT 443  
Db 361 CCTGTCTTATGATGATTAATAATTAAGTAACTGAAATTAAGTAAAGTAAAGTAAAGT 420  
QY 444 CACTCAGCTCAGACTGCCCATGGGACTCTACTACAGCTTCTCCAGAACTTAAACAC 503  
Db 421 CACTCAGCTCAGACTGCCCATGGGACTCTACTACAGCTTCTCCAGAACTTAAACAC 480  
QY 504 GAGAGAAATGTTTGAAGACACAGACACTGTGTGACCTCCATTAATTAAGTAAAGTAA 563  
Db 481 GAGAGAAATGTTTGAAGACACAGACACTGTGTGACCTCCATTAATTAAGTAAAGTAA 540  
QY 564 AATTTCACATGGGCTGATGAATTAAGGACTCTGAGAAAGATAGAACTGATCA 623  
Db 541 AATTTCACATGGGCTGATGAATTAAGGACTCTGAGAAAGATAGAACTGATCA 600  
QY 624 CATTGAGTGGAGTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 683  
Db 601 CATTGAGTGGAGTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
QY 684 TAAATGTAATTCCTGTGAAGAAAAAGAAATTAAGTAAAGTAAAGTAAAGTAAAGTAA 743  
Db 661 TAAATGTAATTCCTGTGAAGAAAAAGAAATTAAGTAAAGTAAAGTAAAGTAAAGTAA 720  
QY 744 CTTGCTCCAGAGAGGTTTGGCAAAATGACAGAGAGAGTCTGCTGAGGAAATAT 803  
Db 721 CTTGCTCCAGAGAGGTTTGGCAAAATGACAGAGAGAGTCTGCTGAGGAAATAT 780  
QY 804 CTACACCATGAGAGAAAGTATATGAGTGAATTAAGTAAAGTAAAGTAAAGTAAAGTAA 863  
Db 781 CTACACCATGAGAGAAAGTATATGAGTGAATTAAGTAAAGTAAAGTAAAGTAAAGTAA 840  
QY 864 CAACAGCCAGAGCCATCTCTGA 885  
Db 841 CAACAGCCAGAGCCATCTCTGA 862

RESULT 5  
US-10-252-131-22  
; Sequence 22, Application US/10252131  
; Publication No. US20030158399A1  
; GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/252,131  
FILING DATE: 20-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-Mar-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-Jun-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-Mar-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 843 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-10-252-131-22  
Query Match 31.1%; Score 843; DB 12; Length 843;  
Best Local Similarity 100.0%; Pred. No. 5.4e-245;  
Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 40 ATGTTTGAAGTCTTACCCCTCAACTGTGCTGCTGCTGCACTACTACTTGCAAGG 99  
DB 1 ATGTTTGAAGTCTTACCCCTCAACTGTGCTGCTGCTGCTGCACTACTACTTGCAAGG 60  
QY 100 TCATTGGAAGATGTTATTAAGTTGAGTTGTTGTAATAATGCTATCTGCTCGCAATTAC 159  
DB 61 TCATTGGAAGATGTTATTAAGTTGAGTTGTTGTAATAATGCTATCTGCTCGCAATTAC 120  
QY 160 ACTCTACCTACATCTGGGACACTTGTGCTTATGCTGTGGGCAAGGAGATTCTGCTTGG 219  
DB 121 ACTCTACCTACATCTGGGACACTTGTGCTTATGCTGTGGGCAAGGAGATTCTGCTTGG 180  
QY 220 TCACAGTGAACAAATGAGTTGCTCAGAACTGATGAAAGAAATGATCAATCAGAAATCC 279  
DB 181 TCACAGTGAACAAATGAGTTGCTCAGAACTGATGAAAGAAATGATCAATCAGAAATCC 240  
QY 280 AGCAGATACCACTAAAGGCGCATCTCAACAAAGAGATGTCTCTGATCTATAAGAT 339  
DB 241 AGCAGATACCACTAAAGGCGCATCTCAACAAAGAGATGTCTCTGATCTATAAGAT 300  
QY 340 GTGACTCTGATGACCTAGGAGCTTAATCTCTGCAAGATCAAGTTCCCTGCTTATGAAT 399  
DB 301 GTGACTCTGATGACCTAGGAGCTTAATCTCTGCAAGATCAAGTTCCCTGCTTATGAAT 360

QY 400 GATATAAATTAAGACTGAATTAAGATCAATCAAGCAGCCAGCTCAGCTCAGACT 459  
DB 361 GATATAAATTAAGACTGAATTAAGATCAATCAAGCAGCCAGCTCAGCTCAGACT 420  
QY 460 GCCCATGGGACTCTACTACAGCTTCTCCAGAACCTTAACCCGAGAGAAATGTTCA 519  
DB 421 GCCCATGGGACTCTACTACAGCTTCTCCAGAACCTTAACCCGAGAGAAATGTTCA 480  
QY 520 GAGACACAGACACTGCTGATACCTCCATATTAACATGGAACAAATTTCCCATGGCT 579  
DB 481 GAGACACAGACACTGCTGATACCTCCATATTAACATGGAACAAATTTCCCATGGCT 540  
QY 580 GATGAATTAAGACTCTGAGAAACGATCAAGCTGATTCACATTTGAGAGAGTC 639  
DB 541 GATGAATTAAGACTCTGAGAAACGATCAAGCTGATTCACATTTGAGAGAGTC 600  
QY 640 TCTGCTGGGTTGACCTTGCACTTATCATTTGGTGTCTTAATCTTAATGTTATTCCTGT 699  
DB 601 TCTGCTGGGTTGACCTTGCACTTATCATTTGGTGTCTTAATCTTAATGTTATTCCTGT 660  
QY 700 AAGAAAAGAAATTTGAGTTGAGCTTATTAACATGCGCCAACTTGCTCCAGAGGG 759  
DB 661 AAGAAAAGAAATTTGAGTTGAGCTTATTAACATGCGCCAACTTGCTCCAGAGGG 720  
QY 760 TTGGCAATGACAGAGCAGTCAAGATGCTCTGAGAAAATATCTACACATCGAGAG 819  
DB 721 TTGGCAATGACAGAGCAGTCAAGATGCTCTGAGAAAATATCTACACATCGAGAG 780  
QY 820 AACGTATATGAAGTGAAGAAATTCAAATAGTACTGCTACGTCAACGAGCAGAGCCA 879  
DB 781 AACGTATATGAAGTGAAGAAATTCAAATAGTACTGCTACGTCAACGAGCAGAGCCA 840  
QY 880 TCC 882  
DB 841 TCC 843

RESULT 6  
US-10-004-633-22  
Sequence 22, Application US/10004633  
Publication No. US20030069196A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
APPLICANT: Lloyd, Clare M.  
APPLICANT: McCarthy, Sean A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
DIAGNOSIS OF IMMUNE DISORDERS  
FILE REFERENCE: 7853-125  
CURRENT APPLICATION NUMBER: US/10/004,633  
PRIOR FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583  
PRIOR FILING DATE: EARLIER FILING DATE: 1996-03-01  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/487,748  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/398,633  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 22  
LENGTH: 843  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: primer  
US-10-004-633-22

Query Match 31.1%; Score 843; DB 14; Length 843;  
Best Local Similarity 100.0%; Pred. No. 5.4e-245;  
Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 40 ATGTTTGAAGTCTTACCCCTCAACTGTGCTGCTGCTGCACTACTACTTGCAAGG 99

Db 1 AAGTTTTCAGGCTTACCTGCACTGTGTCTGCTGCTGCACTACTACTTGCAGG 60  
QY 100 TCATTGGAAGATGTTATTAAGTTGAGTTGTTAAAAATGCTATCTGCTCCAGTTAC 159  
Db 61 TCATTGGAAGATGTTATTAAGTTGAGTTGTTAAAAATGCTATCTGCTCCAGTTAC 120  
QY 160 ACTCTACCTACATCTGTGGGACACTTGTGCTTATGTGCTGGGGCAAGGATTTCTCTTG 219  
Db 121 ACTCTACCTACATCTGTGGGACACTTGTGCTTATGTGCTGGGGCAAGGATTTCTCTTG 180  
QY 220 TCACAGTGTACCAATGATGTGCTCAGAACTGATGAAAGAAATGACATATCAGAAATCC 279  
Db 181 TCACAGTGTACCAATGATGTGCTCAGAACTGATGAAAGAAATGACATATCAGAAATCC 240  
QY 280 AGCAGATACCAAGCTAAAGGGCGATCTCAACAAAGAGATGTCTCTGATCATMAAGAT 339  
Db 241 AGCAGATACCAAGCTAAAGGGCGATCTCAACAAAGAGATGTCTCTGATCATMAAGAT 300  
QY 340 GTGACTCTGATATCAATGAGGACTTGTGCTGAGATATAGTTCCCTGGTCTTATGAT 399  
Db 301 GTGACTCTGATATCAATGAGGACTTGTGCTGAGATATAGTTCCCTGGTCTTATGAT 360  
QY 400 GATTAATAATTAAGATCTGAATATAGACATCAAGACAGCCAGGTCACTCCAGCTCAGACT 459  
Db 361 GATTAATAATTAAGATCTGAATATAGACATCAAGACAGCCAGGTCACTCCAGCTCAGACT 420  
QY 460 GCCCATGAGGAGCTCTACTACAGCTTCTCCAAAGACCTTAAACGAGAGAGAAATGATTC 519  
Db 421 GCCCATGAGGAGCTCTACTACAGCTTCTCCAAAGACCTTAAACGAGAGAGAAATGATTC 480  
QY 520 GAGACACAGACACTGTGAGCCCTCCATTAACAAAGAGAGAAATGATTC 579  
Db 481 GAGACACAGACACTGTGAGCCCTCCATTAACAAAGAGAGAAATGATTC 540  
QY 580 GATGAATAATTAAGATCTGTGAGAAAGAGATCAGAACTGCTTCCACTTGGAGTGGAGTC 639  
Db 541 GATGAATAATTAAGATCTGTGAGAAAGAGATCAGAACTGCTTCCACTTGGAGTGGAGTC 600  
QY 640 TCTGCTGGGTTGACCTGCGCACTTATCATGTTGTTTAAATGATATTCCTGT 699  
Db 601 TCTGCTGGGTTGACCTGCGCACTTATCATGTTGTTTAAATGATATTCCTGT 660  
QY 700 AAGAAAAGAGATTAATGAGTTGAGTGTGAGCTTATCACTGCGCACTTGGCTCCAGAGAG 759  
Db 661 AAGAAAAGAGATTAATGAGTTGAGTGTGAGCTTATCACTGCGCACTTGGCTCCAGAGAG 720  
QY 760 TTGGCAAAATGACAGAGCAGTCAAGATTCGCTGAGGAAATATCTTACACATCGAGAG 819  
Db 721 TTGGCAAAATGACAGAGCAGTCAAGATTCGCTGAGGAAATATCTTACACATCGAGAG 780  
QY 820 AACGTATATGAGTGTGAGAAATTAATGAGTACTACTGCTTACGCTCAACAGCCAGACCA 879  
Db 781 AACGTATATGAGTGTGAGAAATTAATGAGTACTACTGCTTACGCTCAACAGCCAGACCA 840  
QY 880 TCC 882  
Db 841 TCC 843

## RESULT 7

US-10-252-131-23  
Sequence 0, Application US/10252131  
Publication No. US20030158399A1  
GENERAL INFORMATION:  
APPLICANT: Levanson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York

STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/252,131  
FILING DATE: 20-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: Human 200 gene nucleotide  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2236 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 42...944  
SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
US-10-252-131-23  
Query Match 18.4%; Score 499.8; DB 12; Length 2236;  
Best Local Similarity 68.5%; Pred. No. 2.4e-140;  
Matches 842; Conservative 1; Mismatches 333; Indels 53; Gaps 9;  
QY 17 CGAATTCCTCCCAAGTACTGATGTTTCAAGTCTTACCTCACTGTGCTGCTGC 76  
Db 19 CTGACTTTTCTTCTGCAAGCTCCAGTTTTCACATCTTCCCTTGTGCTGTCTGCTGC 78  
QY 77 TCGTGAATCTACTACTTGAAGTCACTTGAAGATGTTATTAAGTTGAGTTGTTGTTAA 136  
Db 79 TCGTGAATCTACTACTTGAAGTCACTTGAAGATGTTATTAAGTTGAGTTGTTGTTAA 138  
QY 137 ATGCTATCTGCTCCAGTACTACCTCTTACCTTACCTTGTGAGCACTTGTGCTATGCT 196  
Db 139 ATGCTATCTGCTCCAGTACTACCTCTTACCTTGTGAGCACTTGTGCTATGCT 198  
QY 197 GGGGCAAGGATTTCTGCTTGTGTCACAGTGTACCAATGAGTTGCTCAGAACTGATGAA 256  
Db 199 GGGGCAAGGATTTCTGCTTGTGTCACAGTGTACCAATGAGTTGCTCAGAACTGATGAA 258  
QY 257 GAAATGTGACATATGAGAAATCAGAGATACAGTAAAGGCGATCCAAAGAGAG 316  
Db 259 GGAATGTGATATATGAGATC--CAGATCTGCTTAAATGAGGATTTCCGAAAGAG 315  
QY 317 ATGTGCTCTGATCAATAAAGATGTGACTCTGATGATGATGATGATGATGATGATGAT 376  
Db 316 ATGTGCTCTGATCAATAAAGATGTGACTCTGATGATGATGATGATGATGATGATGAT 375  
QY 377 TACAGTTCCCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 436  
Db 376 TCCAAATCCAGGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435

437 CCAGGCTACCTCCAGCTGAGCTGCCCATGGGAGCTACTACAGCTTCTCCAGAACCC 496  
 436 CCAGGCTACCTCCAGCTGAGCTGCCCATGGGAGCTACTACAGCTTCTCCAGAACCC 495  
 497 TAACCAAGGAGGAAATGG---TTGAGAGACACACACTGCTGACCTCTCCATATAACA 553  
 496 TTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 555  
 554 ATGGAACCAAAATTTCCAC-----ATGGGCTGATGAA 586  
 556 ATCTAACCAAAATTTCCACATTTGGCCCATGAGTTACGGGAGCTTGAATTTGGCCCATG 615  
 587 TTAAGGAGCTCTGGAGAAACGATCAGACTGCTATCCATGAGTGGAGTCTGCTG 646  
 616 TACGGGAGCTCTGGAGCAACCATCAGAAATAGGCATCTACATGGAGCAGGAGATCTGCTG 675  
 647 GGTGACCTCTGGAGCTTATCATGCTGCTTATCTTAAATGATTCCTGTAAGAAA 706  
 676 GGGCTGCTCTGGAGCTTATCTTAAATGATTCCTGTAAGAAA 735  
 707 AGAAGTTACGAGTTGAGCTTATTAACCTGAGCACTGCTCCAGAGGAGGTTGGCA 766  
 736 AGAAGTTACGAGTTGAGCTTATTAACCTGAGCACTGCTCCAGAGGAGGTTGGCA 795  
 767 ATGAGGAGCAGTCAAGATTCGCTCTGAGGAAATATCTACACATCAGAGGAGGAT 826  
 796 ATGAGGAGCAGTCAAGATTCGCTCTGAGGAAATATCTACACATCAGAGGAGGAT 855  
 827 ATGAGGAGCAGTCAAGATTCGCTCTGAGGAAATATCTACACATCAGAGGAGGAT 886  
 856 ATGAGGAGCAGTCAAGATTCGCTCTGAGGAAATATCTACACATCAGAGGAGGAT 915  
 887 GGGCTGCTCTGGAGCTTATTAACCTGAGCACTGCTCCAGAGGAGGTTGGCA 935  
 916 AACCTTGGGTTGCTGCTTTCGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 975  
 936 TATTTCCCTTTTGGAGAAATCTATGATGATGATGATGATGATGATGATGATGATGAT 995  
 976 GTTGTGCTTTTTCGAGAACTATGATGATGATGATGATGATGATGATGATGATGATGAT 1033  
 996 CCAGGAGCAGTCAAGATTCGCTCTGAGGAAATATCTACACATCAGAGGAGGAT 1055  
 1034 CCAGGAGCAGTCAAGATTCGCTCTGAGGAAATATCTACACATCAGAGGAGGAT 1092  
 1056 GAGTGTACTCATGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1112  
 1093 AACCTGAGCCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1152  
 1113 GATCTCCAGCTCAGAGACT--TCATTCATGCTGTTGAGGAGTCACTGCTGCTT--CATA 1169  
 1153 AGTTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1212  
 1170 CATTGGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198  
 1213 CACTAGGAATCTTGTGAGCTGGGCTCTG 1241

RESULT 8

US-10-004-633-23  
 ; Sequence 23, Application US/10004633  
 ; Publication No. US20030069196A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Levinson, Douglas A.  
 ; APPLICANT: Lloyd, Clare M.  
 ; APPLICANT: McCarthy, Sean A.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
 ; FILE REFERENCE: 7853-125  
 ; CURRENT APPLICATION NUMBER: US/10/004,633  
 ; PRIOR FILING DATE: 2001-12-04  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583

17 CGAATTCCTCCCAAGTACTCAATGTTTGAAGCTTACCTCAAGTGTCTGCTGC 76  
 19 CTGACTTTCTTCTGCAAGCTTCATGTTTCAATCTTCCCTTGAAGTGTCTGCTGC 78  
 77 TCTGCAACTACTACTTTCGCAAGTCTATGGAAGATGTTTAAAGTTGAGTTGTTAA 136  
 79 TCTGCTCTCTACTACTTTCGCAAGTCTATGGAAGATGTTTAAAGTTGAGTTGTTAA 138  
 137 ATGCTATCTGCTCTGCAAGTCTATGGAAGATGTTTAAAGTTGAGTTGTTAA 196  
 139 ATGCTATCTGCTCTGCAAGTCTATGGAAGATGTTTAAAGTTGAGTTGTTAA 198  
 197 GGGGCAAGGAGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256  
 199 GGGGCAAGGAGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 258  
 257 GAAATGTCATATTCAGAAATTCAGAGTATCAGCTTAAAGGAGCTTCAACAAAGAG 316  
 259 GGAATGTCATATTCAGAAATTCAGAGTATCAGCTTAAAGGAGCTTCAACAAAGAG 315  
 317 ATGCTCTCTGATATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 376  
 316 ATGCTCTCTGATATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 375  
 377 TACAGTCTCTGATATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 436  
 376 TCAAAATTCAGAGTATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 435  
 437 CCAGGAGTCTCAGAGTCTGAGCTGCTGAGGAGCTTACTACAGCTTCTCCAGAACCC 496  
 436 CCAGGAGTCTCAGAGTCTGAGCTGCTGAGGAGCTTACTACAGCTTCTCCAGAACCC 495  
 497 TAACCAAGGAGGAAATGG---TTGAGAGACACACACTGCTGACCTCTCCATATAACA 553  
 496 TTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 555  
 554 ATGGAACCAAAATTTCCAC-----ATGGGCTGATGAA 586  
 556 ATCTAACCAAAATTTCCACATTTGGCCCATGAGTTACGGGAGCTTGAATTTGGCCCATG 615  
 587 TTAAGGAGCTCTGGAGAAACGATCAGACTGCTATCCATGAGTGGAGTCTGCTGCTG 646  
 616 TACGGGAGCTCTGGAGCAACCATCAGAAATAGGCATCTACATGGAGCAGGAGATCTGCTG 675  
 647 GGTGACCTCTGGAGCTTATCATGCTGCTTATCTTAAATGATTCCTGTAAGAAA 706  
 676 GGGCTGCTCTGGAGCTTATCTTAAATGATTCCTGTAAGAAA 735  
 707 AGAAGTTACGAGTTGAGCTTATTAACCTGAGCACTGCTCCAGAGGAGGTTGGCA 766  
 736 AGAAGTTACGAGTTGAGCTTATTAACCTGAGCACTGCTCCAGAGGAGGTTGGCA 795  
 767 ATGAGGAGCAGTCAAGATTCGCTCTGAGGAAATATCTACACATCAGAGGAGGATGAT 826

Query Match 18.4%; Score 499.8; DB 14; Length 2236;  
 Best Local Similarity 68.5%; Pred. No. 2.4e-140;  
 Matches 842; Conservative 1; Mismatches 333; Indels 53; Gaps



```

Db      796 ATGCAATGACGAGGGAATTCCTCTCAGAGAAAACATCTATACCATTTGAAGAGAT 855
Qy      827 ATGAATGAGAAATTCMAATGATGACTGCTACGTCACACAGCCACCATCTGAC 886
Db      856 ATGAATGAGAGAGCCCAATGATATATATGCTATGTCAGACAGGACCAACCCCTCAC 915
Qy      887 CGCCTCTGACCTGCCACTTTTAAAGC-----TGCCTTCATTTTCTGACTTTGG 935
Db      916 AACCTTTGGGTTCCTGCTTTCAGATGTCATGATCCAAACCACTTATTTTGAAGCTTGT 975
Qy      936 TATTTCCCTTTTGTGAAAATATATGATGATGATGATGATGATGATGATGATGATGAT 995
Db      976 GTTTTGTCTTTTTCAGAAATATATGATGATGATGATGATGATGATGATGATGATGAT 1033
Qy      996 CCAAGCCACTGAGAAAAGATTCAGTTTCTGGGGATTAATTAATCAACAGGGGATTC 1055
Db      1034 CCACTGCTATGAGAGAGAGTTTCCCATTTTTCAGAAATTA-TGACTCATGGGAATTG 1092
Qy      1056 GACTGTACTCATGCTACATTTGAATGCTCCATTTTATCCCTGAGTTTCAGGG--ATCG 1112
Db      1093 AACCTGGACCTGACATGATTAACAGGACATGATGCTCTGTATTTTAAACCAACAG 1152
Qy      1113 GATCTCCACTCCAGAGACT-TCATATGATGATGATGATGATGATGATGATGATGATGAT 1169
Db      1153 AGTTACCAACCAAGAGACTGTATATGATGATGATGATGATGATGATGATGATGATGAT 1212
Qy      1170 CATTAGGAATGTTAGTGTGATGCTTTG 1198
Db      1213 CACTAGGAATTTTGAAGCTGGGGTCTCTG 1241

```

```

RESULT 9
US-10-188-012-30
; Sequence 30, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyf, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE OF INVENTION: Use Thereof
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: H. sapiens
US-10-188-012-30

```

```

Query Match      17.9%; Score 485.8; DB 14; Length 1116;
Best Local Similarity 70.0%; Pred. No. 2,7e-136;
Matches 759; Conservative 1; Mismatches 278; Indels 47; Gaps 6;
Qy      17 CGGATTTCCCTCCCAAGTACTCATGTTTCAGGTCTTACCTTCATGTCGTCTGTC 76
Db      35 CTGACTTTTCTTTCGAGGCTCATGTTTTCACATCTTCCCTTGAAGTGTCTGTC 94
Qy      77 TCGTCACTACTACTTTCAGAGTCAATGTAAGATGTTTAAAGTGAAGTGTGTA 136
Db      95 TCGTCTGCTACTACTTTCAGAGTCTTCAGAAAGTGAATACAGGCGGAGTGTGTA 154
Qy      137 ATGCTATCTGCTGCTGAGTTACCTCTACCTTACATCTGAGCACTTGTGCTATGTC 196
Db      155 ATGCTATCTGCTGCTGCTTACACCCACCGCCGAGGAACTGTGCGCGCTGCT 214

```

```

Qy      197 GGGGCAAGGGAATTCCTGCTGCTGCTACAGTATACCAATGATGCTCAGAACTGATGAA 256
Db      215 GGGGCAAGGGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 274
Qy      257 GAAATGTGACATATCAGAAATCCAGAGATACAGCTTAAAGGCGATCTCAACAAAGAG 316
Db      275 GGAATGTGATTAATTTGGAATC---CAGATACGCTTAAATGGGATTTCCGAAAGAG 331
Qy      317 ATGTGCTCTGATCATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 376
Db      332 ATGTGCTCTGACATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 391
Qy      377 TACAGTTCCCTGCTCTTATGAATGAATTAATTAATTAATTAATTAATTAATTAATTAAT 436
Db      392 TCCAAATCCAGCAATTAATGAATGAATTAATTAATTAATTAATTAATTAATTAATTAAT 451
Qy      437 CCAAGGTCACTCAGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496
Db      452 CCAAGGTCACTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 511
Qy      497 TAAACACGAGAGAAATG---TTCAGAGACACAGCACTGCTGCTGCTGCTGCTGCTGCT 553
Db      512 TTAACACAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 571
Qy      554 ATGGAACAAATTTTCAC-----ATGGGCTGATGAA 586
Db      572 ATCTAACCAAAATTCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 631
Qy      587 TTAAGACTCTGAGAGAAACAGATCAGATCTATCCATTTGAGAGAGTCTGCTGCTGCT 646
Db      632 TACGGACTCTGAGAGAAACAGATCAGATCTATCCATTTGAGAGAGTCTGCTGCTGCT 691
Qy      647 GGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
Db      692 GGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 751
Qy      707 AGAATTTACAGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 766
Db      752 AGAATTTACAGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 811
Qy      767 ATGCAAGACAGTCAAGATTCGCTGAGAGAAATATCTACACATGAGAGAGATGAT 826
Db      812 ATGCAATACAGAGAGATTCGCTGAGAGAAATATCTACACATGAGAGAGATGAT 871
Qy      827 ATGAATGAGAAATTCATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 886
Db      872 ATGAATGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 931
Qy      887 CGCCTCTGACCTGCCACTTTTAAAGC-----TGCCTTCATTTCTGACTTTGG 935
Db      932 AACCTTTGGGTTCCTGCTTTCAGATGTCATGATCCAAACCACTTATTTTGAAGCTTGT 99
Qy      936 TATTTCCCTTTTGTGAAAATATGATGATGATGATGATGATGATGATGATGATGATGAT 995
Db      992 GTTTTGTCTTTTTCAGAAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1049
Qy      996 CCAAGCCACTGAGAAAAGATTCAGTTTCTGGGGATTAATTAATCAACAGGGGATTC 1055
Db      1050 CCACTGCTATGAGAGAGAGTTTCCATTTTTCAGAAATTA-TGACTCATGAGGAATTG 1108
Qy      1056 GACTG 1060
Db      1109 AACTG 1113

```

```

RESULT 10
US-10-188-012-32
; Sequence 32, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyf, Rosemarie

```



APPLICANT: Kuchroo, Vijay  
 APPLICANT: Freeman, Gordon J.  
 TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
 TITLE OF INVENTION: Use thereof  
 FILE REFERENCE: STAN-235  
 CURRENT APPLICATION NUMBER: US/10/188,012  
 CURRENT FILING DATE: 2002-07-01  
 PRIOR APPLICATION NUMBER: 60/302,344  
 PRIOR FILING DATE: 2001-06-29  
 NUMBER OF SEQ ID NOS: 36  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 32  
 LENGTH: 1116  
 TYPE: DNA  
 ORGANISM: H. sapiens  
 US-10-188-012-32

Query Match 17.9%; Score 485.8; DB 14; Length 1116;  
 Best Local Similarity 70.0%; Pred. No. 2.7e-136;  
 Matches 759; Conservative 1; Mismatches 278; Indels 47; Gaps 6;

17 CGGATTTCCCTCCCAAGTACTGTTTCAAGTCTTACCTCAACTGTGCTGCTG 76  
 35 CTGACTTTCTTCTGCAAGCTCAGTTTTCACATCTTCTTGACTGTGCTGCTG 94  
 77 TCTGCAACTACTTACTTTCAGAGTCAATGGAAGATGTTATAGGTGAGTTGTA 136  
 95 TCTCTGCTGTAATCTTACAGAGTCTTCAAGAGTGAATACAGAGGAGTCCGTA 154  
 137 ATGCTATCTGCTGCTGCACTTCACTTCACTTCTGGAACACTTGTGCTATGCT 196  
 155 ATGCTATCTGCTGCTGCTTCTACACCCGCGCCAGGAACTCGTCCGCTGCT 214  
 197 GGGGGAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256  
 215 GGGGGAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 274  
 257 GAAATGTGATCATGAAATGATGATGATGATGATGATGATGATGATGATGAT 316  
 275 GGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 331  
 317 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 376  
 332 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391  
 377 TACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436  
 392 TCCAAATCCGAGCATATGATGATGATGATGATGATGATGATGATGATGATG 451  
 437 CCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 496  
 452 CCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 511  
 497 TACACGAGGAGAAATG---TTGAGAGACAGACAGTGTGACCTCTCAATAATACA 553  
 512 TTACACGAGGAGAAATG---TTGAGAGACAGACAGTGTGACCTCTCAATAATACA 571  
 554 ATGGAACAAATTTCCAC-----ATGGGCTGATGAA 586  
 572 ATCTAACACAAATTTCCAC-----ATGGGCTGATGAA 631  
 587 TTAAGGACTCTGAGAGAAAGATCAGAACTGATCCATGAGAGGAGTCTGCTGCTG 646  
 632 TACGGAGCTCTGAGAGAAAGATCAGAACTGATCCATGAGAGGAGTCTGCTGCTG 691  
 647 GGTGACCTGAGAGAAAGATCAGAACTGATCCATGAGAGGAGTCTGCTGCTGCTG 706  
 692 GGTGACCTGAGAGAAAGATCAGAACTGATCCATGAGAGGAGTCTGCTGCTGCTG 751  
 707 AAGAGTATGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 766  
 752 AAGAGTATGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 811

767 ATGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 826  
 812 ATGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 871  
 827 ATGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 886  
 872 ATGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 931  
 887 GGTGACCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 935  
 932 AACCTTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 991  
 936 TATTTCCCTTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 995  
 992 GTTTTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1049  
 996 CCACGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1055  
 1050 CCACGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1100  
 1056 GACTG 1060  
 1109 AACTG 1113

RESULT 11  
 US-10-252-131-37  
 Sequence 37, Application US/10252131  
 Publication No. US20030158399A1  
 GENERAL INFORMATION:  
 APPLICANT: Levinson, Douglas A.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
 NUMBER OF SEQUENCES: 37  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036/2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/252,131  
 FILING DATE: 20-SEP-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/609,583  
 FILING DATE: 01-MAR-1996  
 APPLICATION NUMBER: US 08/487,748  
 FILING DATE: 07-JUN-1995  
 APPLICATION NUMBER: US 08/398,633  
 FILING DATE: 03-MAR-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cortuzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7853-048  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-790-9090  
 TELEFAX: 212-869-8864  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 37:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 903 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA  
 SEQUENCE DESCRIPTION: SEQ ID NO: 37:





```
/ Sequence 36, Application US/10414378
/ Publication No. US20030165981A1
/ GENERAL INFORMATION:
/ APPLICANT: FUKUSHIMA, DAIRICHI
/ APPLICANT: SHIBAYAMA, SHIRO
/ APPLICANT: TADA, HIDEAKI
/ TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF
/ FILE OF INVENTION: THE BOTH
/ FILE REFERENCE: 058769
/ CURRENT APPLICATION NUMBER: US/10/414,378
/ PRIOR FILING DATE: 2003-04-16
/ PRIOR APPLICATION NUMBER: US/09/529,063
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: PCT/JP98/04514
/ PRIOR FILING DATE: 1998-10-06
/ PRIOR APPLICATION NUMBER: JP 9-274674
/ NUMBER OF SEQ ID NOS: 117
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 36
/ LENGTH: 2091
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: sig_peptide
/ LOCATION: (53)..(115)
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: (116)..(607)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (53)..(607)
US-10-414-378-36

Query Match      12.1%; Score 328.8; DB 12; Length 2091;
Best Local Similarity 62.4%; Pred. No. 2.2e-88;
Matches 750; Conservative 1; Mismatches 298; Indels 153; Gaps 9;

QY      17 CGGATTTCCCTCCCAAGTACTCATGTTTTCAGGCTTTACCTCACTGTCCTGCTGC 76
DB      30 CTGACTTTTCTTCGCAAGCTCCATGTTTTCACATCTTCCCTTGACTGTCCTGCTGC 89
QY      77 TGTGCACTACTACTTGAAGGTCATTTGAAAGTGTAAAGTTAGGTTGGTGTAA 136
DB      90 TGTGCTGCTACTACTTACAGAGTCTCAAGAGTGAATACAGACGAGAGTGGTACA 149
QY      137 ATGCTATCTGCTGCTGAGTTACACTCTACCTACCTGAGACACTTGTGCTATGTGCT 196
DB      150 ATGCTATCTGCTGCTGCTGCTTACACCCGAGCCGCCAGGGAACCTGTGCTGCTGCT 209
QY      197 GGGGCAAGGATTTCTGCTTGTGTCACAGTGTACCAATGAGTTGCTCAGAACTGATGAA 256
DB      210 GGGGCAAGGAGGCTGCTGCTGCTGTTGAATGTGCAACGTTGCTCAGACCTGATGAA 269
QY      257 GAAATGACATATCAGAAATCCAGAGATACAGCTAAAGGGGATCTCAACAAAGAG 316
DB      270 GGAATGATATTTTGAATC---CAGATATGCTGCTAAATGGGATTTCCGCAAGGAG 326
QY      317 ATGTGTCTGTATCATTAAGATGTACTGTGATGACATGAGGACCTACTGTGTCAGGA 376
DB      327 ATGTGTCTGTATCATTAAGATGTACTGTGATGACATGAGGATCTACTGTGCTGCGGA 386
QY      377 TACAGTTTCTGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 436
DB      387 TCCAAATCCAGGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 446
QY      437 CCAAGGCTACTCAGCTCAGCTGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496
DB      447 CCAAGGCTACTCAGCTCAGCTGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 506
QY      497 TAAACACGAGAGAAATG---TTACAGACACACAGACACTGTGTGACCTTCATATATACA 553
DB      507 TTACACACGAGGAGCATGGCCGACAGACAGACACAGACACTGTGTGAGGAGCTTCCTGATATTA 566
```

```
QY      554 ATGGAACAAAATTTCCACATGCGCTGATGAATTAAGACTGTGAGAAACGATCAGAA 613
DB      567 ATCTAAC----- 574
QY      614 CTGCTATCCACATTGAGTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
DB      575 ----- 574
QY      674 TCTTAATCTTAATGATATTCCTGTAAGAAAAGATTATGACTTTGAGCTTATTA 733
DB      575 -----GTAATTCATAGCAAAAGAAAGATACAGATTTAAGCTCATCT 619
QY      734 CACTGGCCAACTTGCTCTCAGAGAGGTTGGCAATTCAGAGACAGTCTGCTGCTG 793
DB      620 CTTTGGCCAACTTCCTCTCCTCAGAGTGGCAATTCAGATGAGAGGAAATTCCTCAG 679
QY      794 AGGAAATATCTACACCATCGAGAGAGATATGAGTGAAGTGAATTCAAATGACTACT 853
DB      680 AAGAAAACATCTATACCATTTGAAGAGAGATATGAGTGAAGTGAAGTGAAGTGAAGTGA 739
QY      854 ACTGCTACGTCACACAGCAGACCCATCTGACCCGCTCTGACCTGACCTTTAAAGC 913
DB      740 ATTGCTATGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 799
QY      914 -----TGCCTTCACTTTCTGACTTTGATTTTCCCTTTTGTGAAAATATGTA 962
DB      800 CATGATTCACACCATTTATTTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 859
QY      963 TATGCTACTTGGCACTCATGAGGATTCAGACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1022
DB      860 TGTGTACCTGACTGCTGTT---TTGAGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 917
QY      1023 TTTTCTGGGATATTAATTAATCAACAGGAGATTCAGCTGTAATCATGCTATGTAATG 1082
DB      918 TTTTCAAGAGATTA---TGACTCAGATGGAATTAAGCTGAGACCTGACATGAATTAACA 976
QY      1083 CTCATTTTATCCCTGAGTTTCAAGG---ATGGAATCTCCCATCCAGAGACT---TCAATC 1138
DB      977 GGCATGCTATGCTGCTGCTGATTTTAACCAACAGAGTATCCCAACCAAGAGACTGTTAATC 1036
QY      1139 ATGCTGTTGAAGCTCAGCTGCTGCTTTC---ATACATTAAGAAATGTTAGTGTGCTT 1196
DB      1037 ATGATGTTAGAGCTCAAAACGGCTTTTATATACCTAGAAATTTCTTGAAGTGGGCTTC 1096
QY      1197 TG 1198
DB      1097 TG 1098

RESULT 15
US-09-529-063-35
/ Sequence 35, Application US/09529063
/ Patent No. US20020102542A1
/ GENERAL INFORMATION:
/ APPLICANT: FUKUSHIMA, DAIRICHI
/ APPLICANT: SHIBAYAMA, SHIRO
/ APPLICANT: TADA, HIDEAKI
/ TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF
/ FILE OF INVENTION: THE BOTH
/ FILE REFERENCE: 058769
/ CURRENT APPLICATION NUMBER: US/09/529,063
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: PCT/JP98/04514
/ PRIOR FILING DATE: 1998-10-06
/ PRIOR APPLICATION NUMBER: JP 9-274674
/ NUMBER OF SEQ ID NOS: 117
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 35
/ LENGTH: 555
/ TYPE: DNA
/ ORGANISM: Homo sapiens
```

US-09-529-063-35

Query Match	10.7%	Score 289;	DB 10;	Length 555;
Best Local Similarity	73.7%;	Pred. No. 1.1e-76;		
Matches 396;	Conservative 0;	Mismatches 135;	Indels 6;	Gaps 2;

Qy	40	ATGTTTTCAGGTTCTTACCCCTCAACATGTGTCGTCGTCGTCGCAACTATCTATCTTGGAAAG	99
Ds	1	ATGTTTTCACATCTTCCCTTTGACGTGTGTCCTGCTGCTGCTGCTGCTATCTATCTTACAAAG	60
Qy	100	TCATTGGAAAGATGGTTATTAAGTTGAGTTGGTTAAATAATGCTATCTGCCCTGCAGTTAC	159
Ds	61	TCTCTCAGAAGTGGAAATACAGAGGGAGTCGGTTCAGATGCTATCTGCCCTGCTTAC	120
Qy	160	ACTCTACTTACATCTGCGGACATCTTGCTGCTATGTGCTGCGGCAAGGATTCGTCTTGG	219
Ds	121	ACCCCAAGCGCCCCAGAGGAACCTCGGCCGCTGCTGCTGCGGCAAGGAGCCTGTCTGTG	180
Qy	220	TCACAGTGTACCAATAGTTGCTCAGAACTGATGAAAGAAATGTGACATATCAGAAATCC	279
Ds	181	TTTGAATGTGGCAACGTGTGCTCAGAGACTGATGAAAGGATGTGAATTATTGGACATC	239
Qy	280	AGCAGATACCGCTAAAGGCGCATCTCAACAAGAGATGTGTCTGTGATCTATAAAT	339
Ds	240	--CAGATACGCGCTAAATGGGGATTTCCGCCAAAGAGATGTGCTCTGACCATAGAAAT	297
Qy	340	GTGACTCTGGATACCATGAGGGAACCTACTGCTCAGAGATACAGTTCCCTGATCTATGAAT	399
Ds	298	GTGACTCTAGCAACAGGTGGATCTTACTGCTGCGGATCCAAATCCCAAGCATATGAT	357
Qy	400	GATPAAAAATTTAGAACTGAAATTGACATCAAGAAGCCAAAGTCACTCCAGCTCAGACT	459
Ds	358	GATGAAAAATTTTAACCTGAAGTTGTGCATCAAAACAGCAAGGTCAACCTTGACACGACT	417
Qy	460	GCCCATGGGGAATCTTACTACAGTTTCTCAAGAACCTTACCACGAGAGAAATG---T	516
Ds	418	CTGCAGAGACATCTCACTCAGGCTTTTCCAAAGATGCTTACCACCAAGGGGACATGGGCCCA	477
Qy	517	TCAGAGACACAGACACTGGTGACCCCTCCATATTAATCAATGGAACAAAAATTTCCACA	573
Ds	478	GCAAGACACAGACACTGGGAGACCTCCTGATATTAATCTTACACGATATCTCATATA	534

Search completed: November 22, 2003, 05:38:22  
Job time : 886.58 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 23:34:40 ; Search time 290.42 Seconds  
(without alignments)  
10161.809 Million cell updates/sec

Title: US-10-004-633-37

Perfect score: 903  
Sequence: 1 atgttccacatctccctt.....gtgtcgttgcattgcacga 903

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 216961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	903	100.0	903	US-10-252-131-37	Sequence 37, Appl
2	903	100.0	903	US-10-004-633-37	Sequence 37, Appl
3	903	100.0	1116	US-10-188-012-10	Sequence 30, Appl
4	903	100.0	2236	US-10-252-131-23	Sequence 0, Appl
5	903	100.0	2236	US-10-004-633-23	Sequence 23, Appl
6	901.4	99.8	1116	US-10-188-012-32	Sequence 32, Appl
7	585	64.8	2091	US-09-529-063-36	Sequence 36, Appl
8	585	64.8	2091	US-10-414-378-36	Sequence 36, Appl
9	522.8	57.9	555	US-09-529-063-35	Sequence 35, Appl
10	522.8	57.9	555	US-10-414-378-35	Sequence 35, Appl
11	456	50.5	2710	US-10-252-131-8	Sequence 8, Appl
12	456	50.5	2710	US-10-004-633-8	Sequence 8, Appl
13	456	50.5	2725	US-10-188-012-10	Sequence 10, Appl
14	445	50.4	529	US-10-040-739-415	Sequence 415, Appl
15	442.6	49.0	843	US-10-252-131-22	Sequence 22, Appl
16	442.6	49.0	843	US-10-004-633-22	Sequence 22, Appl

17	440.8	48.8	862	14	US-10-188-012-12	Sequence 12, Appl
18	396.6	43.9	1203	10	US-09-966-546-21	Sequence 21, Appl
19	396.6	43.9	1203	10	US-09-966-545-21	Sequence 21, Appl
20	396.6	43.9	1203	11	US-09-965-212-21	Sequence 21, Appl
21	396.6	43.9	1203	15	US-10-189-940-21	Sequence 14, Appl
22	333.6	36.9	375	15	US-10-027-632-258124	Sequence 258124, Appl
23	82.8	9.2	582	12	US-10-027-632-258124	Sequence 258124, Appl
24	82.8	9.2	582	13	US-10-027-632-258125	Sequence 258125, Appl
25	81.6	9.0	582	12	US-10-027-632-258126	Sequence 258126, Appl
26	81.6	9.0	582	13	US-10-027-632-258126	Sequence 258126, Appl
27	81.6	9.0	582	13	US-10-027-632-258126	Sequence 24, Appl
28	81.6	9.0	582	13	US-10-027-632-258126	Sequence 18, Appl
29	72.6	8.0	1079	14	US-10-188-012-24	Sequence 20, Appl
30	72.6	8.0	1080	14	US-10-188-012-20	Sequence 26, Appl
31	72.6	8.0	1080	14	US-10-188-012-20	Sequence 22, Appl
32	72.6	8.0	1095	14	US-10-188-012-26	Sequence 28, Appl
33	72.6	8.0	1099	14	US-10-188-012-28	Sequence 2, Appl
34	72.6	8.0	1099	14	US-10-188-012-28	Sequence 14, Appl
35	66.4	7.4	918	14	US-10-188-012-2	Sequence 16, Appl
36	66.4	7.4	1032	14	US-10-188-012-14	Sequence 4, Appl
37	66.4	7.4	1032	14	US-10-188-012-16	Sequence 20, Appl
38	65.4	7.2	849	14	US-10-188-012-4	Sequence 20, Appl
39	65.2	7.2	571	9	US-09-739-907-20	Sequence 732, Appl
40	64.8	7.2	398	10	US-09-960-352-7732	Sequence 1063, Appl
41	64.8	7.2	418	10	US-09-960-352-1063	Sequence 34, Appl
42	64.4	7.1	1156	14	US-10-188-012-34	Sequence 36, Appl
43	64.4	7.1	1156	14	US-10-188-012-36	Sequence 68, Appl
44	64.4	7.1	1282	11	US-09-813-153-68	Sequence 59, Appl
45	64.4	7.1	1380	14	US-10-037-270-59	

#### ALIGNMENTS

RESULT 1  
US-10-252-131-37  
Sequence 37, Application US/10252131  
Publication No. US20030158399A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/252,131  
FILING DATE: 20-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US/08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US/08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Cornuzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864

```

?       TELEX: 66141 PENNIE
?       INFORMATION FOR SEQ ID NO: 37:
?       SEQUENCE CHARACTERISTICS:
?           LENGTH: 903 base pairs
?           TYPE: nucleic acid
?           STRANDEDNESS: single
?           TOPOLOGY: linear
?       MOLECULE TYPE: DNA
?       SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-10-252-131-37

```

[illegible]

QY	901	CCA	903
Db	901	CCA	903
QY	901	CCA	903
Db	901	CCA	903

RESULT 2

US-10-004-633-37

Sequence 37, Application US/10004633

Publication No. US20030069196A1

GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.

APPLICANT: Lloyd, Clare M.

APPLICANT: McCarthy, Sean A.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND

FILE REFERENCE: 7853-125

CURRENT APPLICATION NUMBER: US/10/004,633

CURRENT FILING DATE: 2001-12-04

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986

PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583

PRIOR FILING DATE: EARLIER FILING DATE: 1996-03-01

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/487,748

PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/398,633

NUMBER OF SEQ ID NOS: 49

SOFTWARE: PatentIn Ver. 2.0

SEO ID NO 37

LENGTH: 903

TYPE: DNA

ORGANISM: Homo sapiens

US-10-004-633-37

Query Match . 100.0%; Score 903; DB 14; Length 903;

Best Local Similarity 100.0%; Pred. No. 4,1e-292;

Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY	1	ANGTTTTCACATCTTCCCTTTGACCTGTGCTGCTGCTGCTGCTACTACTTACAGG	60
Db	1	ANGTTTTCACATCTTCCCTTTGACCTGTGCTGCTGCTGCTACTACTTACAGG <td>60</td>	60
QY	61	TCCTCAGAGTGGAAATACAGACGCGAGGTGGTCAGAAATGCTATCTGCCCTTCTTAC <td>120</td>	120
Db	61	TCCTCAGAGTGGAAATACAGACGCGAGGTGGTCAGAAATGCTATCTGCCCTTCTTAC <td>120</td>	120
QY	121	ACCCGACCGCCGCCGAGGAACTCGTGGCCCTCTGCTGGGGCAAAAGAGCTTCCCTG <td>180</td>	180
Db	121	ACCCGACCGCCGCCGAGGAACTCGTGGCCCTCTGCTGGGGCAAAAGAGCTTCCCTG <td>180</td>	180
QY	181	TTTGAATGTGGCAAGTGGTGTCTAGAGCTATGAAAGGGATGTGAATTATTTGAACATCC <td>240</td>	240
Db	181	TTTGAATGTGGCAAGTGGTGTCTAGAGCTATGAAAGGGATGTGAATTATTTGAACATCC <td>240</td>	240
QY	241	AGATCTGCTAAATGGGGATTTCCGCAAGAGAGATGTGTCCTGACCATAGAGATGTG <td>300</td>	300
Db	241	AGATCTGCTAAATGGGGATTTCCGCAAGAGAGATGTGTCCTGACCATAGAGATGTG <td>300</td>	300
QY	301	ACTCTAGCAGACAGTGGGATCTACACGCTCGCGGATCCAAATTCGACGGCATATGATGAT <td>360</td>	360
Db	301	ACTCTAGCAGACAGTGGGATCTACACGCTCGCGGATCCAAATTCGACGGCATATGATGAT <td>360</td>	360
QY	361	GAAAAATTAACTGAAGTGGTCACTAAACACAGCAAGGTCACCCCTGGACCGACTGTG <td>420</td>	420
Db	361	GAAAAATTAACTGAAGTGGTCACTAAACACAGCAAGGTCACCCCTGGACCGACTGTG <td>420</td>	420
QY	421	CAGAGAGCTTCACTGACGCTTTTCCAGAGATGCTTACACAGGGGACATGGCCACGCA <td>480</td>	480
Db	421	CAGAGAGCTTCACTGACGCTTTTCCAGAGATGCTTACACAGGGGACATGGCCACGCA <td>480</td>	480
QY	481	GAGACACAGACCTGGGAGCCTCTGTATATTAATCTAAACAAATATCCATTGGCC <td>540</td>	540























**THIS PAGE BLANK (USPTO)**



**TITLE**  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

**JOURNAL**  
MEDLINE  
PUBMED  
Contact: Yoshihide Hayashizaki  
22534683  
12466851

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 677)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Miyakido, I., Otsu, N., Saito, R., Suzuki, H., Yamahara, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bull, C., Hume, D.A., Quackenbush, J., Schmitt, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brady, D., Brusic, V., Chochia, C., Corbett, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontus, J.U., Qi, D., Ring, Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yamaoka, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Murata, M., Nomura, K., Numaaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takekida, Y., Waki, K., Watanabe, A., Muramatsu, M., and Hayashizaki, Y. Direct  
Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

JOURNAL  
MEDLINE  
COMMENT

CONTACT: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
URL: <http://genome.gsc.riken.go.jp/>  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numaaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takekida, Y., Waki, K., Watanabe, A., Muramatsu, M., and Hayashizaki, Y. Direct  
Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES  
source

Location/Qualifiers  
1..677  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NOD"  
/db\_xref="taxon:10090"  
/clone="F630119G06"  
/cell\_type="NOD-derived CD11c +ve dendritic cells"  
/clone\_id="RIKEN full-length enriched, NOD-derived CD11c +ve dendritic cells"  
BASE COUNT 199 a 151 c 155 g 167 t 5 others  
ORIGIN

Query Match 20.4% Score 552.8: DB 14; Length 677;  
Best Local Similarity 97.1%; Pred. No. 4.2e-75;  
Matches 560; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

253 GAAAGAAATGTGACATATCAGAAATCCAGAGATACCGATTAAGGGCATCTCAACAA 312  
|||  
101 GAAATGCTTATGNGTTGAGAAATCCAGAGATACCGATTAAGGGCATCTCAACAA 160  
|||  
313 GGAATGTCTCTGATCATTAAGATGTACTCTGATGACCATGGAGCTTACTGCTGC 372  
|||  
161 GGAGACGTCTCTGATCATTAAGATGTACTCTGATGACCATGGAGCTTACTGCTGC 220  
|||  
373 AGAATACGTTCCCTGCTTATGATGATTAAGATTAAGAACTGAATTAAGATCAAA 432  
|||  
221 AGAATACGTTCCCTGCTTATGATGATTAAGATTAAGAACTGAATTAAGATCAAA 280  
|||  
433 GCAGCAAGTCTACTCCAGCTCAGACTGCCATGGAGGAGCTACTACAGCTTCTCAAGA 492  
|||  
281 GCAGCAAGTCTACTCCAGCTCAGACTGCCATGGAGGAGCTACTACAGCTTCTCAAGA 340  
|||  
493 ACCCTAACCGAGAGAAATGTTTCAAGACACAGACCTGTGACCTCCATTAATAC 552  
|||  
341 ACCCTAACCGAGAGAAATGTTTCAAGACACAGACCTGTGACCTCCATTAATAC 400  
|||  
553 AATGGAACAAATTTCCATGAGGCTGATGAATTAAGAGCTCGAGAAACGATCGA 612  
|||  
401 AATGGAACAAATTTCCATGAGGCTGATGAATTAAGAGCTCGAGAAACGATCGA 460  
|||  
613 ACTGCTATCCATTTGAGTGGAGTCTCTGCTGGTGTGACCTGGCACTTATCTTGGT 672  
|||  
461 ACTGCTATCCATTTGAGTGGAGTCTCTGCTGGTGTGACCTGGCACTTATCTTGGT 520  
|||  
673 GTCTTAATCTTAAATGTTATCTCTGTAAGAAAGAAAGTTATCGAGTTGACCTTAT 732  
|||  
521 GTCTTAATCTTAAATGTTATCTCTGTAAGAAAGAAAGTTATCGAGTTGACCTTAT 580  
|||  
733 ACACGGCCACCTTCCCTCCAGAGGCTTGGCAATGACAGACGATCGATTCCT 792  
|||  
581 ACACGGCCACCTTCCCTCCAGAGGCTTGGCAATGACAGACGATCGATTCCT 640  
|||  
793 GAGGAATATCTACCATCGAGAGAACTATATG 829  
|||  
641 GAGGAATATCTACCATCGAGAGAACTATATG 677  
|||

RESULT 3  
BE68651/c 491 bp mRNA linear EST 11-SEP-2000  
LOCUS BE68651/c  
DEFINITION uw03c07.x1 Soares mouse 3kbMS Mus musculus cDNA clone IMAGE:3415596  
3', mRNA sequence.  
ACCESSION BE68651  
VERSION BE68651.1 GI:10074275  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 491)  
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),



Oy	377	TACGATCCCGGCTTATGAATGAATGAATAAATAATTAGAACTGAATTAAGACATCAAGCGAG	436
Db	430	TCCAATCCCAAGGCATATGAATGAATGAATAAATAATTAACTGAAGTTGTATCAAAACGAG	489
Oy	437	CCAAGTCACTCCAGCTCAGACTGCGCCATGGGGACTCTAACAAGCTTCTCCAGAACC	496
Db	490	CCAAGTCAACCCCTGCACCGACTCTGAGAGAGACTCACTGACGGCTTTCCAAAGATGC	549
Oy	497	TAAACACGAGAGAAATG---TTGAGAGACACAGACACTGTGACCCCTCCATAATACA	553
Db	550	TTACCAACAGGGGACATGGGCCACAGAGACACAGACACTGGGAGACCTCCGATATATA	609
Oy	554	ATGGAACAAAATTTCCAC-----ATGGCTGATGAAA	586
Db	610	ATCTAACAACAATATCCAATTCGGCCAAATGATTAACGGGACTCTGAATTTGGCCAAATGACT	669
Oy	587	TTAAGACTCTGAGAGAAACATCAGACTCTCTATCCACATTTGAGTGGAGTCTGTGCTG	646
Db	670	TACGGACTCTGAGAGCAACCATCAGAAATAGGCATCTACATCGAGACAGGATCTGTGCTG	729
Oy	647	GGTGACCCCTGGACCTTATCAATGGTGTCTTAAATCCTTAAATGTAATTCCTGTAGAAAA	706
Db	730	GGCTGGCTCTGGCTCTTATCTTGCGGCGCTTAAATTTCAATGTAATTCATACGAAAG	789
Oy	707	AGAAGTATAGAGTTTAGCCTATTATCACTGGCCAACTTGCTCCAGAGGGTTGGCAA	766
Db	790	AGAAATACAGAAATTAAGCCTCATCTTTGGCGAACCTCCCTCCAGATTGGCAA	849
Oy	767	ATGCAGAGCAGTCAGAGATTGCGCTGTGAGGAAATATCTACACCATCGAGAGACGTAT	826
Db	850	ATGCAGATGACAGAGGAAATGCGTCAGAGAAACATCTATACCATTTGAAGAAAGTAT	909
Oy	827	ATGAAGTGAAGAAATTCAAATGAGTACTACTAGCTCAACGACGACAGCCATCTCTGAC	886
Db	910	ATGAAGTGAAGAGAACCCAAATGAGATTAATTTGATATGACAGACGAGGCAACCTCTAC	969
Oy	887	CGCCTCTGAGCTGCCACTTTTAAAGC-----TGCCCTTATTTCTGACTTTGGT	936
Db	970	AACCTTTGGGGTGTGCGTTTGGCAGMGCCAMARATCCAAACCATTAATTTTGAOCTTGATG	1029
Oy	937	ATTTCCTCTTTGTGAGAAACTATGATGATATGCACTTGGCAACCTCATTTGGAGGTCTGCAC	996
Db	1030	TTTTTGTCTTTTTCGAAATATATAGCTGTGTACCTG---ACTGGTTTGGRGKTYTGT	1085
Oy	997	CACAGCACACTGAGAAAAGAGTTCCAGTTTCTGSGGATAATTA	1040
Db	1086	CMCTGCTATMTGRTGMAAGTTTTCCHTTTCARAAAAAAMATTCMM	1129

RESULT 5  
 BX526231  
 ID BX526231 standard; RNA; EST; 476 BP.  
 XX BX526231;  
 AC BX526231.1  
 SV BX526231.1  
 XX  
 XX 27-MAY-2003 (Rel. 75, Created)  
 DT 27-MAY-2003 (Rel. 75, Last updated, Version 1)  
 XX  
 DE RZPD Mus musculus cDNA clone IMAGE998J53740 = IMAGE:1478438 5' EST.  
 XX  
 KW EST; expressed sequence tag.  
 XX  
 OS Mus musculus (house mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 XX  
 RN [1]  
 RP 1-476  
 RA Heil O., Ebert L., Neubert P., Peters M., Radelof U., Schneider D.,  
 RA Korn B.;  
 RT ;

RJ	Submitted (28-MAY-2003) to the EMBL/GenBank/DDBJ databases.
RJ	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Im Neuenheimer
RJ	Feld 580, D-69120 Heidelberg, Germany
XX	
CC	RZPD; IMAGp98G153740.
CC	RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
CC	Mouse Unigeneset - RZPD2 (RZPDLIB No.981)
CC	http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=981
CC	Contact: Ina Rottis
CC	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
CC	Heubenerweg 6, D-14059 Berlin, Germany
CC	Tel.: +49 30 32639 101
CC	Fax: +49 30 32639 111
CC	www.rzpd.de
CC	This clone is available royalty-free from RZPD;
CC	contact RZPD (clone@rzpd.de) for further information.
-CC	Seq primer: T7, Primer sequence: TAATACGACTCACTATAGGG
XX	
PH	Key
FT	Location/Qualifiers
FT	source
FT	1. .476
FT	/db_xref=taxon:10090"
FT	/note="1st strand cDNA was primed with a Not I - Oligo(dT)
FT	primer [15' TGTTACCATTGAAATGAATGAGCGGCCGGTATCTTTTTTTT
FT	TTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to
FT	EcoRI adaptors 5'-AATTCGCACGAG-3' and 5'-CTGTGCTCG-3'
FT	(pharmacia), digested with NotI and cloned into the NotI
FT	and EcoRI sites of the pTZ19-Pac vector. Library is
FT	normalized. Library was constructed by Beno Soares and M.
FT	Patiama Bonaldo."
FT	/organism="Mus musculus"
FT	/clone="IMAGp98G153740"
FT	/clone_1id="Soares NMPu pregnant uterus"
FT	/dev_stage="adult"
FT	/lab_host="DH10B"
XX	
SQ	Sequence 476 BP; 135 A; 124 C; 113 G; 104 T; 0 other;
	Query Match 16.5%; Score 446; DB 4; Length 476;
	Best Local Similarity 95.9%; Pred. No. 1.1e-58;
	Matches 468; Conservative 4; Mismatches 4; Indels 12; Gaps 1;
QY	1539 GAATCTATATTGTAATCGTAGACCCCACAATGAAAAAGCTAGGCTGTGTAAGACATGCTTGT 1598
DB	1 GAATCTATATTGTAATCGTAGACCCCACAATGAAAAAGCTAGGCTGTGTAAGACATGCTTGT 60
QY	1599 AGACTCAAAGAATGAGAGGTTAAAGGACAAACAATCCCCGGGGCTTGCGTGCAGTCAGC 165
DB	61 AGACTCAAAGAATGAGAGGTTAAAGGACAAACAATCCCCGGGGCTTGCGTGCAGTCAGC 120
QY	1659 TTAGCTAGAGTGTGAGTTCCTCAATGCCAACAGATCCCTGTCCTAMAGTAAGATGACRTG 1718
DB	121 TTAGCTAGAGTGTGAGTTCCTCAATGCCAACAGATCCCTGTCCTAMAGTAAGATGACRTG 180
QY	1719 AGTATCTGGCGCATGTCCATGGGGGTTGTCTCTCTCTCAGAAAGACATGCATGWC 1778
DB	181 AGTATCTGGCGCATGTCCATGGGGGTTGTCTCTCTCTCAGAAAGACATGCATGAAA 240
QY	1779 CCTGCATGMAATGA 1838
DB	241 CCTG-----CACACACACACACACACACACACACACACACACACATGMAATGA 288
QY	1839 AGGTTCTCTGTGCTGCTGCTACTCTCTATPAACATGTATCTCTACAGACTCTCCTCTGC 1898
DB	289 AGGTTCTCTGTGCTGCTGCTACTCTCTATPAACATGTATCTCTACAGACTCTCCTCTGC 348
QY	1899 CTCTGTTAAGCATGAGTGGGAGCATGGACAGCAGTCCAGTAATTTATTTACAGACTCA 1958
DB	349 CTCTGTTAAGCATGAGTGGGAGCATGGACAGCAGTCCAGTAATTTATTTACAGACTCA 408
QY	1959 GAAAGCTGGACAGAAAGCTGGAGAAATTCAGAGACATCTGGCCCAACATCTGCCAGACTCT 2018
DB	409 GAAAGCTGGACAGAAAGCTGGAGAAATTCAGAGACATCTGGCCCAACATCTGCCAGACTCT 468

Qy	2019	TCTTACAC	2026
Db	469	TCTTACAC	476

LOCUS	BB750705	440 bp	mRNA	EST 19-OCT-2001
DEFINITION	BB7507075	Riken full-length enriched, pooled tissues, cerebellum,		
	etc. Mus	musculus cDNA clone GI3020116 3',	mRNA sequence.	
ACCESSION	BB750705			
VERSION	BB750705.1	GI:16154941		
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			

**TITLE** RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al 2001)

## FEATURES

```
/organism="Mus musculus"  
/mol_type="mRNA"  
/strain="C57BL/6J"  
/db_xref="taxon:10090"  
/clone=GJ30201N16"  
/clone_lib="Riken full-length enriched, pooled tissues,  
cerebellum, etc."  
/note="Pooled tissues; (tissue_type=cerebellum,  
dev_stage=16 days neonate, sex=mixed),  
(tissue_type=cerebellum, dev_stage=0 day neonate,  
sex=mixed), (tissue_type=hippocampus, dev_stage=adult,  
sex=male), (tissue_type=whole body, dev_stage=9 days
```

BASE COUNT	embryo, sex=mixed)	(tissue_type=lung, dev_stage=13 days
ORIGIN	embryo, sex=mixed)"	
127 a	83 c	101 g
		129 t





BASE COUNT 106 a 97 c 86 g 85 t  
 ORIGIN

Query Match 12.8%; Score 345.6; DB 9; Length 374;  
 Best Local Similarity 95.1%; Pred. No. 2.8e-43;  
 Matches 367; Conservative 4; Mismatches 3; Indels 12; Gaps 1;

QY 1540 AATCTATATTTGATTCGTAGACCCCACTGAAAAGCTGAGCTGTGTAGACATGCTTTGA 1599  
 DB 1 AATCTATATTTGATTCGTAGACCCCACTGAAAAGCTGAGCTGTGTAGACATGCTTTGA 60  
 QY 1600 GACTCAGAGATGAGAGGTAAAGGCAACAAGATCCCGGGGCTTGCTGACATGACCT 1659  
 DB 61 GACTCAGAGATGAGAGGTAAAGGCAACAAGATCCCGGGGCTTGCTGACATGACCT 120  
 QY 1660 TAGCTAGATGCTGAGTTCAGTCCAGTCCAGAGATCCCTGTCTCAGAGATGAGCTGA 1719  
 DB 121 TAGCTAGATGCTGAGTTCAGTCCAGTCCAGAGATCCCTGTCTCAGAGATGAGCTGA 180  
 QY 1720 GTATCTGGCGCATGCTCATGCGGGGTTGCTCTCTCTCAGAGAGACATGACATGACC 1779  
 DB 181 GTATCTGGCGCATGCTCATGCGGGGTTGCTCTCTCTCAGAGAGACATGACATGACC 240  
 QY 1780 CTGCATGAAATGAA 1839  
 DB 241 CTG-----CACACACACACACACACACACACACACACACACACATGAAATGAA 288  
 QY 1840 GGTCTCTCTGTGCTGCTGCTGCTCTCTTAATCAATGATCTCTCAAGACTCTCTGACC 1899  
 DB 289 GGTCTCTCTGTGCTGCTGCTGCTCTCTTAATCAATGATCTCTCAAGACTCTCTGACC 348  
 QY 1900 TCTGTTAAGCATGATGCTGGAGCATG 1925  
 DB 349 TCTGTTAAGCATGATGCTGGAGCATG 374

RESULT 9 354 bp mRNA linear EST 30-SEP-1998  
 LOCUS A1152783  
 DEFINITION uc91b01.x1 Soares\_NMPu Mus musculus cDNA clone IMAGE:1478185 5',  
 mRNA sequence.  
 A1152783  
 ACCESSION A1152783.1 GI:3681252  
 VERSION  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 354)  
 Maria, M., Hillier, J., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
 Scheinberg, K., Stepec, M., Tan, F., Underwood, K., Moore, B.,  
 Treising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and  
 Waterston, R.  
 TITLE The WashU-HMI Mouse EST Project  
 JOURNAL Unpublished  
 COMMENT Contact: Maria M/Mouse EST Project  
 WashU-HMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu  
 This clone is available royalty-free through LIND; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 WGI:926561  
 Seq primer: -28ml3 rev2 ET from Amersham  
 High quality sequence stop: 330.  
 Location/Qualifiers

FEATURES  
 source 1..354  
 /organism="Mus musculus"  
 /mol\_type="mRNA"

/db\_xref="taxon:10090"  
 /clone="IMAGE:1478185"  
 /sex="female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares\_NMPu"  
 /note="Organ: uterus; Vector: pUT73D-Pac (Pharmacia) with  
 a modified polylinker; 1st strand cDNA was prepared from  
 pregnant mouse uterus, and was then primed with a Not I -  
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
 RI adaptors (Pharmacia), digested with Not I and cloned  
 into the Not I and Eco RI sites of the modified pUT73  
 vector. Library is normalized. Library was constructed by  
 Bento Soares and M. Fatima Bonaldo."

BASE COUNT 102 a 87 c 91 g 74 t  
 ORIGIN

Query Match 12.0%; Score 324; DB 9; Length 354;  
 Best Local Similarity 94.5%; Pred. No. 5.9e-40;  
 Matches 346; Conservative 4; Mismatches 4; Indels 12; Gaps 1;

QY 1492 CTGAGATGAGGCTGCGACATGAGTGTGTGTAACAAGCAGGAATCTAATTTG 1551  
 DB 1 CTGAGATGAGGCTGCGACATGAGTGTGTGTAACAAGCAGGAATCTAATTTG 60  
 QY 1552 ATGCTGACCCCAATGAAAGCTAGGCTGTGATGAGATGCTTGTAGATCAAGAT 1611  
 DB 61 ATGCTGACCCCAATGAAAGCTAGGCTGTGATGAGATGCTTGTAGATCAAGAT 120  
 QY 1612 GGAGAGTAAAGGACACACACATCCCGGGGCTTGCTGACATGACCTTAGCTAGTGC 1671  
 DB 121 GGAGAGTAAAGGACACACACATCCCGGGGCTTGCTGACATGACCTTAGCTAGTGC 180  
 QY 1672 TGAATTCACATGACACACAGTCTCTCAGAAAGATGAGTATCTGCGCA 1731  
 DB 181 TGAATTCACATGACACACAGTCTCTCAGAAAGATGAGTATCTGCGCA 240  
 QY 1732 TGTCCATGGGGGTGCTCTCTCTCAGAAAGATGAGTATCTGCGCA 1791  
 DB 241 TGTCCATGGGGGTGCTCTCTCTCAGAAAGATGAGTATCTGCGCA 291  
 QY 1792 ACACACACACACACACACACACACACACACACACATGAAATGAGTCTCTGT 1851  
 DB 292 ---CACACACACACACACACACACACACACACACACATGAAATGAGTCTCTGT 348  
 QY 1852 GCCTGC 1857  
 DB 349 GCCTGC 354

RESULT 10 583 bp mRNA linear EST 11-OCT-2002  
 LOCUS B0789118  
 DEFINITION 1040b02.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:6128595  
 3' similar to TR:054947 054947 KIDNEY INJURY MOLECULE-1 PRECURSOR  
 mRNA sequence.

ACCESSION B0789118  
 VERSION B0789118.1 GI:23838264  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 583)  
 Melton, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K.,  
 Lemishka, I., Seacore, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
 Hillier, L., Maria, M., Pape, D., Wyllie, T., Martin, J., Blistein, A.,  
 Schmitt, A., Treising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas  
 M., Gibbons, M., McCann, R., Cole, R., Tsagarisvilli, R., Williams, T.,  
 Jackson, Y., and Bowers, Y.

TITLE Endocrine Pancreas Consortium  
 JOURNAL Unpublished  
 COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue



```

Db      735 AACCCAGCGCGCCAGGGAACCTGCTGCGCTGCGGGGAAAGAGCCTGCTCTG 676
Qy      218 GGTACACAGTTCACCAATGAGTTGCTGAACTGATGAAAGAAATGACATATCAGAAAT 277
Db      675 TGTNTGAAATGTGGCAACGTGTGCTCAGACATGAAAGGAGTGTAAATTTATGGACAT 616
Qy      278 CCAGCAGATACCAAGCTTAAAGGCGATCTCAACAAAGAGATGTCTCTCATATTAAGA 337
Db      615 C---CAGATACGTGCTTAAATGGGGATTTCCGAAAGAGATGTCTCTCAACATAGAGA 559
Qy      338 ATGTGACTCTGATGACATGACCTGCTGCTGACAGATACAGTTCCCTGCTTATGA 397
Db      558 ATGTGACTCTAGACAGACAGTGGATCTACTGCTGCTCCGATCCAAATCCAGGCAATATGA 499
Qy      398 ATGATATAAATTTAGAACTGAATTTAGACATCAAAAGCAGCAAGGTCACTCCAGCTCAGA 457
Db      498 ATGATGAAATTTAATCTGAAAGTTGATCACTAAACAGACAGAG-----ACA 452
Qy      458 CTGCCCCTGGGACTCTACTACAGCTTCTCCAGAACCCCTAACCCAGAGAGAAATGTT 517
Db      451 CAGACACTGGGGAGCT------CCCTGATATTAATCTAA 418
Qy      518 CAGAGACACAGACAGCTGATGACCTCCATATTAACATGGAACAAATTTCCACATGGG 577
Db      417 CACAAATATCCACATTTG-----CCATGAGTTACGGGACTCTAGATGG 373
Qy      578 CTGATGAAATTTAGAGACTCTGAGAAACGATCAGAACTGCTTACCATTTGAGTGGAG 637
Db      372 CCATGACTTACCGGACTCTGAGCAACCATCAGAAATGAGCATCTACATGGAGCAGGGA 313
Qy      638 TCTCTGTGGGTTGACCCCTGCGCACTTATCATGTGCTTAACTTAAATGATTCCT 697
Db      312 TCTGTGTGGGCTGCTGCTGCTGCTTATCTTACGCGCTTAAATTTAAATGATTCCT 253
Qy      698 GTAGAGAAAGAAATTTATGAGTTTGAAGCTTATTAACATGCGCAACTTCTCCAGAG 757
Db      252 ATGCAAAAGAGATGATGAAATTTAGCCTCATCTTTGGCAACCTCCCTCCAG 193
Qy      758 GGTGGCAATTCAGAGAGCAGTCAAGATTGCTCTGAGAAATATTTACACATGAGG 817
Db      192 GATGGCAATTCAGAGAGTTCAGAGAGGAAATTCCTCAGAAAGAAACATTTATCATTGAG 133
Qy      818 AGAACGATATGAGTGAAGTGAATTTCAATAGATCTCTACGTACGTCAACAGCAGAGC 877
Db      132 AGAACGATATGAGTGAAGTGAAGCCTCAATAGATTTATGCTATGTCAGAGAGCAGC 73
Qy      878 CATCTGACCGCCTCTGAGCTGCACTTTTAAAGC 913
Db      72 AACCTTCAACAACCTTTGGTTTGTGCTTGCATATGC 37

RESULT 12
CD520418      863 bp      mRNA      linear      EST 06-JUN-2003
LOCUS      AGENCOURT 14354411 NIH MGC 191 Homo sapiens cDNA clone
DEFINITION      IMAGE:30411529 5', mRNA sequence.
ACCESSION      CD520418.1 GI:31452136
VERSION      CD520418
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 863)
AUTHORS      NIH-MGC http://mgs.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-r@mail.nih.gov

```

```

Tissue Procurement: Narayan Bhat
CDNA Library Preparation: CLOUTIER Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.llnl.gov
Plate: NDCM203 row: h column: 02
High quality sequence stop: 509.
Location/Qualifiers
1..863
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30411529"
/issue_type="Pooled"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 191"
/notes="Vector: pDNR-LIB; Site 1: SfiI (ggccatcaggcc);
Site 2: SfiI (ggcgcccgcc); Library is oligo-dT primed
and directionally cloned. PBMC - Peripheral Blood
Mononuclear Cells. RNA was pooled from 3/hour stimulat
with PMA adn Ionomycin. 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CAAGGCGCATTTAGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGAGCGGCCACACAG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC Library."

BASE COUNT      221 a      222 c      203 g      215 t      2 others
ORIGIN
Query Match      10.8%; Score 291.8; DB 14; Length 863;
Best Local Similarity 72.4%; Pred. No. 3,6e-35;
Matches 420; Conservative 0; Mismatches 153; Indels 7; Gaps 3;

Qy      17 CGGATTTCCCTCCCAAGTATCTATTTTCAAGTTTACCTCAACTGTCTCTGCTG 76
Db      36 CTGACTTTTCTCTGCAAGCTCCTGTTTCAATCTTCTTGAAGTCTGCTGCTG 95
Qy      77 TGCTGCACTACTCTGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 136
Db      96 TGCTGCTGCTACTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAG 155
Qy      137 ATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 196
Db      156 ATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 215
Qy      197 GGGGCAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 256
Db      216 GGGGCAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 275
Qy      257 GAAATGTGACATATGAAATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTC 316
Db      276 GGAATGTGATTTTGGACATC---CAGATCTGCTTAAATGGGATTTCCGAAAG 332
Qy      317 ATGTGTCTGTGATCAATAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 376
Db      333 ATGTGTCTGTGACATATGAAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 392
Qy      377 TACAGTTCTGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATG 436
Db      393 TCCAAATCCAGGCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 452
Qy      437 CCAAGGTCACTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 496
Db      453 CCAAGGTCACTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 512
Qy      497 TAAACACGAGAGAAATGTTGAG--AGACACAGACACT-GGTGACCTTCATTAAT 552
Db      513 TTACACACAGGAGACATGCGCCAGCAGAGACACAGACATGCGGAGGAGCTCC 572

```





Db 597 TTACCACCAAGGGACATGGCCCGACGACAGACACAGACTGGGG 640

Search completed: November 22, 2003, 01:46:24  
Job time : 5944.8 secs

**THIS PAGE BLANK (USPTO)**